Tampere University

Miikka Vuorinen

REQUIREMENTS FOR THE ANALYTICAL PROCESS OF DEMAND FORECASTING

Faculty of Engineering and Natural Sciences

Master of Science Thesis

May, 2029

ABSTRACT

Faculty of Engineering and Natural Sciences
Miikka Vuorinen: Requirements for the analytical process of demand forecasting
Master of Science Thesis, 64 pages
Tampere University
Master's Degree Programme in Information and Knowledge Management
Major: Knowledge Management, Information Analytics
May, 2019

Analytics can be viewed from various perspectives such as set of technologies, tools or processes. Building organizations analytical capabilities should be started from basics rather than pursuing unrealistic high-end analytics straight away. Identifying maturity of analytics in an organization creates the baseline and helps to set targets for the organization's analytical capabilities development. Opportunities to utilize analytics are many and rather than focusing on all of these it is advantageous to start building organizations analytical capabilities from a selected use case which is prominent for an organization.

This research was conducted as an embedded case study and targeted to identify analytical process that supports demand forecasting in an organization. To answer this research problem literature review was conducted to establish theoretical framework around subject and to identify analytical process framework for demand forecasting based on the literature. After this, interviews were conducted to identify requirements in an organization for analytics and demand forecasting. Based on theoretical and empirical parts analytical process for demand forecasting was identified for an organization.

Building analytical process for demand forecasting in an international organization with various functions is not a straight forwarded development. Each function has their own requirements for analytics and demand forecasting. Common nominator with these requirements is a business need behind the process. Starting the process from a business need, guarantees that the process is expected to create value for the organization and for unnecessary forecasts are not created. Forecast is a view of the future based on past information. Therefore, it is important to communicate with relevant stakeholders throughout the demand forecasting process to ensure that right information is used as a basis for the forecast.

Keywords: Analytics, analytics process, forecasting demand, data processor

TIIVISTELMÄ

Tekniikan ja luonnontieteiden tiedekunta
Miikka Vuorinen: Kysynnän ennustamisen analytiikkaprosessin vaatimukset
Diplomityö, 64 sivua
Tampereen yliopisto
Tietojohtamisen diplomi-insinöörin tutkimusohjelma
Pääaine: Tiedon ja osaamisen hallinta, Informaatioanalytiikka
Toukokuu, 2019

Analytiikkaa voidaan lähestyä monista näkökulmista kuten siihen liittyvien teknologioiden, työkalujen tai prosessien. Organisaation analytiikkakyvykkyyden rakentaminen on hyvä aloittaa perusteista, sen sijaan että heti lähdettäisiin tavoittelemaan monimutkaisia analytiikkaratkaisuja. Tunnistamalla organisaation analytiikan maturiteettitaso, saadaan selville lähtötilanne ja voidaan tunnistaa analytiikan kehittämisen tavoitteet. Organisaatioissa analytiikan hyödyntämisen kohteita on monia. Sen sijaan että keskityttäisiin kehittämään näitä kaikkia, on kannattavampaa aloittaa kehitys fokusoitumalla valikoituun, organisaatiolle hyödylliseen analytiikan käyttötapauksen kehittämiseen.

Tämä tutkimus toteutettiin sulautettuna tapaustutkimuksena, jonka tavoitteena oli tunnistaa analytiikkaprosessi, joka tukee kysynnän ennustamista organisaatiossa. Jotta tutkimuksessa voitaisiin vastata tähän tutkimusongelmaan, aiheesta tehtiin kirjallisuuskatsaus, jonka avulla muodostettiin aiheen ympärille teoreettinen viitekehys sekä analyyttisen prosessin viitekehys kysynnän ennustamisen kontekstissa. Tämän jälkeen haastattelututkimuksen keinoin selvitettiin organisaation vaatimukset analytiikalla ja kysynnän ennustamiselle. Lopuksi, tutkimuksen teorian ja empirian kautta tunnistettiin organisaatiolle sopiva kysynnän ennustamista tukeva analytiikkaprosessi.

Analytiikkaprosessin rakentaminen kansainvälisessä organisaatiossa, jossa on useita toimintoja, ei ole suoraviivainen kehityshanke. Jokaisella toiminnolla on omat tarpeensa analytiikalle ja kysynnän ennustamiselle. Yhteisenä nimittäjänä voidaan kuitenkin pitää liiketoiminnallista tarvetta prosessin taustalla. Aloittamalla kysynnän ennustamisen analytiikkaprosessi liiketoimintatarpeesta taataan, että prosessin oletetaan tuottavan arvoa organisaatiossa, sekä myöskään turhia ennusteita ei tehdä. Ennuste on näkymä tulevaisuuteen historiallisen informaation valossa. Tästä syystä on tärkeää kommunikoida olennaisten sidosryhmien kanssa, läpi kysynnän ennustamisprosessin, jotta voidaan varmistua ennusteen pohjautuvan oikeaan informaatioon.

Avainsanat: Analytiikka, analytiikkaprosessi, kysynnän ennustaminen, datan käsittelijä

PREFACE

This Master Thesis was done for Robit Plc. Main parts of the research were conducted during July, August and September of 2018. The three-month period during the which the research was done included refinement of the topic, familiarization to the domain, creating empirical framework, conducting interviews, analyzing empirical findings and writing the conclusion chapter. This period was very hectic and included many long days in the library.

I would like to thank all the people from the target organization from the opportunity to conduct research and write this thesis. Secondly, I want to thank Professor Samuli Pekkola for counseling me throughout the research and especially offering insights on how to narrow the research topic. I also want to thank my friends and family who have listened to my complaints about the research, offered valuable insights and somehow believed that I would graduate one day. This has truly helped me to believe in it also.

Tampere, 20.5.2019

Miikka Vuorinen

CONTENTS

LIST OF FIGURES AND TABLES

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Background & need for research

In the information era, data the source of information has become vital in many situations. Organizations are collecting increasing amounts of data which leads to increasing amount of information (Chen et al. 2012). Organizations have been competing on how most actionable and profitable information can be created from this data for decades now (Davenport 2006). Whether analytics is called business intelligence machine learning or even artificial intelligence, all aim to create more actionable right timed information to support decision making (Tyrväinen 2013). Increasing data amounts and interesting business opportunities have created requirements around analytics which has been seen to affect organizations investments in analytics as well as scientific research around the topic (Davenport 2006; Chen et al. 2012; Tyrväinen 2013; Lee et al. 2014).

Davenport (2006) sees analytics as vital part of organization's competitiveness since in globalizing world numerous competitors can offer comparable products or services but analytics can increase the efficiency of business processes. Efficient demand forecasting is one used case for analytics as it aims to offers insights to many issues such as inventory levels, development of sales and supply chain planning (Stevenson 2014). When developing analytics in organization, approach should be considered carefully. Organizations often have existing data that could be refined to actionable information, but without considering the approach to analytics and full potential of analytics is hard to reach (Krumeich et al. 2014).

Opportunities to record and refine data are everywhere. Industry has moved towards smart factories and products which allows optimizing varying from starting points of the manufacturing processes to predictive maintenance of end products customers are using. Various sensors can be placed different machines opening a gateway to record everything machine does. (Lee et al. 2014) This often results in to diverse and complex data and analytical requirements which call for advanced technologies and techniques to get beneficial results (Chen et al. 2012). This requires organization to consider their own analytical maturity and their analytical requirements. If organization's analytical maturity level is low it is beneficial to start building analytical capabilities from basics rather than aiming towards unreachable goals to begin with (Idoine & Herschel 2016), which is also supported by Chen et al. who acknowledge that modern complex analytics require understanding of analytics utilized in the turn of the century as well as up to date information and vision of the field.

This research is done for Robit Plc, hereinafter referenced as Robit or target organization, which is a publicly limited liability company listed in Nasdaq Helsinki specializing in manufacturing, selling and servicing drilling consumables. Their four major market areas are mining, construction, tunneling and well-drilling. Robit is globally operating company with production sites currently in Australia, Finland, South Korea, United Kingdom and United States of America. Company's business can be divided into three strategic business units, Top Hammer, Down the Hole and Digital Services. (Robit Plc)

1.2 Research problem & questions

Purpose of this research is to determine, how analytics and more precisely, demand forecasting should be done in Robit Plc in order to enable demand forecasting. As noted earlier, this research is focused on developing analytics process for the target organization. When considering analytical process which enables demand forecasting, there are various angles from which the research can be approached. In the target organization it was defined that main three operators in analytical process are data, it's processors and technology used in the process. Since target organization does not have systematic analytic operations or analytical process, technological aspects were deemed to be outside of the scope of this research and the analytical process aspects are limited data and processor requirements.

As Robit wants to be able to forecast their demand more efficiently. As noted by Laihonen et al. (2013) for efficient decision-making it is vital that initial information and requirements have been identified. In the case of demand forecasting this means that aspects affecting forecasting need to be considered as well as what is even considered as demand forecasting. In the scope of this research these are considered especially from data requirements and processors point of view.

From above mentioned aspects, research problems main research question can be stated:

1) What kind of analytical process supports demand forecasting?

In order to be able to answer to the main research question following supporting research questions were identified as well

a) What data sources are required to execute analytics?
b) What tasks are needed and what is required from the data processor in order to execute analytical process?

1.3 Research targets and scope

This research is conducted for Robit Plc as target organization and therefore the scope is defined greatly by, they requirements. Robit wants to be able to forecast demand more efficiently. Scope of the research was originally planned to be predicting demand based on Robit's ERP data and various external data sources such as drilling data from mines globally. Maturity of the analytic processes and existing data in Robit did not enable predictive data-analytics and demand forecasting development and therefore focus of this research shifted towards defining efficient analytic processes.

As mentioned in the Chapter 1.2. this research aims to identify factors required for analytical process that supports demand forecasting. Analytics case was specified according to target organization's requirements, but process is considered also from more general business intelligence and analytics point of view in order to avoid biases created by specified analytical task. As this research does not aim to create analytics, technical solutions such as data warehousing and other specific business intelligence tools are outside the scope. This research aims to develop solid foundation for the target organization to start operating systematic and efficient analytics.

From Robit's business point of view this research is defined to Top Hammer business. This definition was since Robit's Lempäälä site was decided to be the one where interviews were conducted, and site's core business is Top Hammer consumables. In addition, ERP system contained most amount of Top Hammer data. From more practical side, interviews and research outcomes were required to be done in the summer holiday season in 12-week time which meant that scope was required to be quite limited.

1.4 Research structure

This research examines research topic from theoretical and empirical point of view. As the topic of this research can be looked from various perspectives theoretical framework aims to establish how these topics are seen in this research as a background for the empirical part of this research. Both theoretical and empirical parts are then used to establish analytical process that supports demand forecasting.

First chapter, the introduction section of this research introduces the reader to the requirement for research on the topic. Research problem and questions are presented as well as targets and scope of the research. Second chapter focuses on the research methodology of this research. Execution of both literature review and empirical research are introduced also.

Third chapter contains theoretical framework of this research. It is divided into two main topics: demand forecasting and business intelligence & analytics. Starting from the basics of the topic's literature review aims to build general understanding for the reader about

them, focusing on the sides which are in scope of this research. In the final part of the theoretical framework, literature's view of analytical framework supporting demand forecasting is introduced as well as how literature sees the research questions. This framework is used as basis for empirical part of this research.

Fourth chapter introduces the conduction of empirical part of this research. Findings of the interviews are introduced and explained to give reader general understanding on how empirical part was done and what were the findings. Last chapter, Discussions and Conclusions results are combined based on literature review and empirical research. Research questions are answered based on these findings. Finally, critical evaluation of this research is done from the researcher's point of view and possible future research targets are evaluated.

2. RESEARCH METHODOLOGY

2.1 Research methodology

Scientific research requires designing and evaluating the methodology of the research. Saunders et al. (2009) present the research onion as one way to design research. Research onion is an idea where researcher is required to consider design of their research as six layers of the onion where they start by peeling the outermost layer and continuing this layer by layer to the center of the onion. Six layers of the onion are from outermost to the innermost:

1. Research philosophy
2. Research approach
3. Research strategies
4. Research choices
5. Time horizons
6. Data collection and analysis

Considering research philosophy, the first layer of the onion, is the starting point when starting to design the research. Importance of considering research philosophy is that it defines how researcher views the world and the problem under research. Four philosophies of the research are positivism, realism, interpretivism and pragmatism. (Saunders et al. 2009) p108. From philosophical point of view, this research falls under interpretivism because of the ontology, epistemology, axiology and data collection techniques utilized. First from ontology point of view, the research is conducted socially and is subjective to the target organization. Secondly, from epistemology acceptable knowledge for this research can be acquired from the literature and more detailed from the target organization. Thirdly, as the research is done for a target organization in which the researcher has been employed, the researcher is a part of the phenomena being researched and the view is therefore subjective. Lastly, relatively small qualitative data collection methods are utilized.

Second layer of the onion is the research approach. Research approaches can either be inductive, deductive or the combination of them in which case research approach is abductive. Difference between inductive and deductive approaches is the role of the theory in the research. Deductive approach focuses on developing a theory and testing this theory where inductive approach focuses on building a theory based on data (Saunders et al. 2009) p124. As this research focuses on building a theoretical framework and modifying this for the target organization based on empirical findings, approach of the research is

inductive. Research also emphasizes qualitative data collection and is not highly concerned about generalization which are typical for inductive approach (Saunders et al. 2009) p127.

First two layers of the research onion define the research on higher level and consider the research philosophies. As noted by Saunders et al. (2009) p.136 next three layers of the onion focus more on the research process and consider how research questions are actually converted into a research project. For this the purpose of the research must be considered. Research purpose can either be exploratory, descriptive or explanatory but research can also have multiple purposes (Saunders et al. 2009) p.139. In the case of this research, the purpose is to define analytics process and seek new insights from the target organization to specify the process for them. This makes this research exploratory.

Third layer in the research onion is the research strategy. As note by Easterby-Smith et al. (2012) many factors need to be considered when defining research strategy. Saunders et al. (2009) second this notion but also include that the choice of the research strategy is mostly defined by the research questions. In the case of this research, case study was chosen to be the strategy for the research since the problem requires empirical research of the subject in the target organization (Robson & McCartan 2016). More specifically, research is single embedded case study. Single case was chosen since the research is conducted for the target organization and as per research questions, it aims to provide analytical framework for the target organization. Embedded case was chosen also due research because of the research questions. The analytical process requires various actors throughout the process and therefore interviewees are needed from various functions of the organization which makes the case embedded.

Choosing the methods for the research is next layer of the research onion. Analytical framework of the research question is derived from the literature and is then modified based on the conducted interviews. Both data collection methods are qualitative which makes the research multi-method qualitative research (Saunders et al. 2009) p.152. Since the research question aims to produce analytical framework to support demand forecasting for the target organization, specific subject is being studied at the current time. This makes the research time horizon, the fifth layer of the onion, cross-sectional (Saunders et al. 2009) p.154.

2.2 Theoretical framework

Theoretical framework of this research was conducted as literature review. This research aimed to produce value also to the target organization and the scope was originally more high-level analytics. Findings of the literature review were presented to the target organization throughout the review process which iterated the research problem and scope further. Original target of creating predictive analytics based on multiple sources was shifted

towards analytical process development since many research showed that it is advantageous to start building analytical capabilities from low-level analytics, see for example (Idoine & Herschel 2016; Wang, G. et al. 2016). As noted by Saunders et al. (2009) refining the research problem and ideas is one of the fundamental reasons to review literature.

Second reason behind the importance of the literature review, as argued by Saunders et al. (2009), is to familiarize the researcher with the subject. Acquiring and critically evaluating literature around analytics and forecasting formed the basis on how this research views the subject and on the other hand how it fits in the research field in general. This literature basis also improved general understanding analytics and forecasting for both the researcher and the target organization.

As noted by Easterby-Smith (2012) there are various sources that can be utilized in literature review such as books, periodicals, theses, government publications and reference work. In this research literature was chosen based on Saunders et. al (2009) division between primary, secondary and tertiary literature sources. Since the scope of this research required information about the fundamentals of analytics and forecasting rather than most cutting-edge research, literature search was focused especially on secondary and tertiary sources. Primary literature sources are often their first releases and have not necessarily gone through publishing process which includes peer reviews and other evaluations (2009). Taking these into account with strict schedule of this research it was justified not to spend time on finding primary research and focus more on published and peer reviewed research.

To obtain literature for the review, several online databases and search engines available at Tampere University of Technology were utilized as well as library of the university. Google Scholar, Scopus and Andor were the three main databases and search engines used online. As argued above gathering of literature started to obtain more knowledge around subjects of this research and to refine the scope of the research. Following list contains keywords used in online searches:

- “Analytics” AND “Manufacturing industry”
- “Analytics” AND “Forecasting” AND “Mining”
- (“Demand plan*” OR “Forecast*”) AND “Mining industry”
- “Predicting demand” AND “Industry”
- “Business Intelligence” AND “Manufactur*”
- (“Business Intelligence” OR “BI”) AND “Process model”
- “Process model” AND “Forecast*” (OR “Demand plan*)

With utilizing above keywords much literature was found around the topics. From these choices some filtering was required. This was done using search engines’ filtering capabilities by focusing on literature that is peer reviewed literature focusing on maximumly

10 years old researches. Choosing of the literature from filtered options was then made based on researcher's evaluation. Title of the source was evaluated and if deemed appropriate and relevant for the scope of this research abstract of the source was read. If the source was then deemed suitable based on abstract the article was read and evaluated whether it was useful in answering the research question of this research.

Several articles and theses read during the literature review contained useful information for this research, but the part was cited from somewhere else. This led to searching of original sources and their evaluations. Original sources were preferred whenever possible.

2.3 Empirical framework

2.3.1 Interviewees

Interviewees were chosen throughout target organizations from organization's business and supporting functions. Goal of the interviews was to support theoretical framework and whole thesis in answering the research questions. As noted also in the introduction this thesis was done as a case study, in which framework for business intelligence and analytics in the context of forecasting was derived from theoretical background and empirical part focuses fixing this framework specifically for the target organization.

Collecting data through interviews allowed identifying the factors in BI&A process which are relevant outside academic world. Empirical results are presented according to themes that rose up during the interviews. Answers of interviewees are compared to each other to evaluate the themes various point of views.

Interviews were originally planned around four main topics: supply chain process, strategic and operational analytics processes, quality of data and sales. These topics acted as basis for purposive sampling of the interviewees. Product portfolio rose as important theme in the first interview, and it was added as fifth theme. Ten interviews were planned but due one last minute cancellation nine were conducted. Roles of the interviewees and themes and dates of interviews have been collected in the Table 1.

Interviewee	Interview date	Role in the company	Theme of the interview
I1	13.9.2018	Production management and purchasing	Supply chain & data quality
I2	14.9.2018	Sales & analytics, sales and operations planning.	Sales & analytics
I3	14.9.2018	Supply chain & financial analytics	Analytics & data quality
I4	17.9.2018	Production management, purchasing and subcontracting	Supply chain & data quality

I5	17.9.2018	Financial controller, reporting and analytics	Analytics & data quality
I6	17.9.2018	ERP manager, development	Product portfolio & data quality
I7	18.9.2018	Sales and offering support, Top Hammer product line	Product portfolio & sales
Cancelled	18.9.2018 / canceled	Overall sales and offering support	Product portfolio & sales
I8	21.9.2018	Sales	Sales & product portfolio
I9	21.9.2018	Supply chain analytics	Analytics & data quality

Table 1. *Interview dates, interviewees and themes of the interviews*

2.3.2 Interview planning and execution

Empirical material for this research was collected with interviews. Since this thesis was done as a case study for target organization, interviewees were all chosen from within the organization's different functions. As the research problem is diverse and in-depth exploration of the subject was required, qualitative approach for empirical framework was chosen (Saunders et al. 2009). Mixture of purposive sampling method and snowball sampling was used to identify the interviewees. Research was done for organizations in which researcher had been working for. This allowed using purposive sampling based on researchers own expertise about the organization. Counseling regarding interviewees were also asked from organizations management to ensure representativeness of the sample. During the interviews each interviewee were also asked to name other persons who should be interviewed for the thesis. This allowed identifying and exploring factors which affect forecasting.

To explore multiple themes around forecasting and analytics, semi-structured interview method was chosen. As suggested by Saunders et al. (2009) p.320 list of questions around forecasting and analytics was composed as the backbone of the interview. Each interview consisted of general questions which were asked from every interviewee and topic specific questions which were depended on the interviewee's role in the organization. Questions of the interviews were based on theses' research questions and what kind of information is aimed to collect through interviews. List of questions used is available in the Appendix A.

As is typical for semi-structured interviews, all interview varied from each other (Saunders et al. 2009). All topics in the interview list were covered in each interview, but order of the questions varied as well as did the length of the interviews. Additional themes were

also explored during the interviews. For example, external factors which affect production and therefore supply capabilities rose up during an interview and questions around external factors were asked also during following interviews,

All interviews were conducted face to face at organization's Lempäälä office. This promoted free flowing conversation during the interview (Saunders et al. 2009). All interviews were recorded with interviewees approval and later transcribed to notes of the interview. Important themes were additionally taken to note during interview to ensure that these were appropriately included also after the transcription of the recordings. Interview recordings were erased after the thesis was completed.

Short presentation of the subject of the thesis was given before the interviews for all interviewees to familiarize and focus them on the subject. Interviews began with question about persons position in the organization and relationship to analytics and general reporting. There the interviews shifted towards forecasting, available data sources and information needs. Theme specific questions were asked throughout the interview depending on the flow of the conversation. In the end each interviewee was also asked whether they think some theme regarding analytics or forecasting was left undebated at this point. If so, this was further examined, and question list was iterated for upcoming interviews. Lastly each interviewee was asked to name other employees who would be relevant candidates for interviews without given information about other interviewees already chosen.

2.3.3 Interview analysis

Analysis of the interviews was done utilizing qualitative methods. Recorded interviews were transcribed to text. Transcription was done by making detailed notes from each recording and combining dealt subjects rather than transcribing all recordings word to word. Alternative methods suggested by Saunders et al. (2009) p 486 to reduce time needed in transcription were also considered but since amount audio to be transcribed was around 400 minutes listening all files and making notes from them was considered to be best way to execute the transcriptions.

Interviews were conducted mainly in Finnish. Only interviews I2 and I10 were conducted English. Language of the interview was chosen based on whether interviewer and interviewee understand each other better in English or Finnish. Since the thesis' language is English, interviews hold in Finnish required translation. This required translating quotations used in the research which in some cases could have affected the tone of the answer. Each quotation used, was translated using dictionary in ambiguous words to remove bias of the researcher.

Since data collected with qualitative methods is often diverse there are little standardized methods available to analyze it (Saunders et al. 2009) p490. Summarizing the data was

used as first tool to simplify created notes. This was by combining all questions and answers to single spreadsheet which allowed identifying main themes discussed during the interviews. Secondly, since the interviews were all categorized under themes mentioned in the chapter 2.3.1, it made sense to use categorization also with the analysis process. Categorization aimed to further understand the relationship of different factors throughout the forecasting process.

Categorization proved quite difficult since the semi-structured nature of the interviews allowed free-flowing conversation about the subject. This led to long answers, during which many questions were touched. Because of this, each interview needed to be analyzed thoroughly and considered as entity.

3. THEORETICAL FRAMEWORK

3.1 Demand forecasting

Organizations need to be able to forecast demand to answer their customers' needs (Armstrong 2001; Feigin 2014; Stevenson 2014). Based on the forecast organization can then plan their expected demand and utilize these plans in many areas such as logistics, warehouse management and production planning but as noted by Stevenson (2014), forecasting can be utilize in every function of the organization. Demand forecasting aims to either plan the future demand or planning the effects of the demand (Stevenson 2014) p.75. Planning the future demand refers to larger, long-term plans of the future such as locations, product portfolio and other strategic decisions. Planning the effects of the demand refers for example to planning inventory levels and other operative decisions. Insufficient demand forecasting can have negative affects through the supply chain of the organization. Delivery shortages can lead to decrease in customer satisfaction, organizations' expenses can increase for example in warehousing costs or supplier relationships may suffer from continually changing orders (Feigin 2014).

Feigin (2014) states that organizations typically forecast at least two types of demand. Firstly, total sales of the organization are forecasted in diverse ways. This forecast is often based on management's view of the market and the set targets of the organization. It is also often done on a longer time horizon. Secondly organizations are interested to know what the predicted sales of the shorter period is such as next week or month. This forecast is done based on historical data and is referred as unbiased forecast (Feigin 2014). Unbiased forecast mirrors, the future based on past data and when determining its accuracy, it is important to consider how accurately used data actually reflects the future (Stevenson 2014).Vital difference of these forecast is in the purpose of the forecast types: one is done based on aim of the organization and the other is based on historical data and other facts. Both are still needed for efficient forecasting but serve different purpose in organization (Feigin 2014).

As the demand forecasting is potential view of the future, forecasting and planning demand is in fact risk management and essentially forecasting should be exercised by all who do business. Even though the importance of forecasting is generally recognized in organizations, it is incorrect to assume that efficient forecasting repot will automatically be answer to all demand planning challenges in organization. (Feigin 2014) Organizations are also required to consider what they should forecast. Generally forecasting should not be done unless there is uncertainty about the future (Armstrong 2001). It is important to understand that forecast is just a tool to be used to enhance view of the business and provide information for decision making.

Before investing in forecasting organization should consider expected return of the investment. Forecasts provide information for decision making and the returns from them are therefore indirect (Stevenson 2014). Since return of investment in case of forecasting comes indirectly from for example reduced safety stocks organization needs be ready to follow the plan based on the improved forecast to get return (Feigin 2014). These expected indirect returns should be identified at early stage of investment planning since management needs to get benefits from the investment.

When organization decides to invest in forecasting the question is how to get most out of the investment? Answering this is no simple task but investing in forecasting should be considered based on the purpose of forecast which should be considered early in the acquisition (Stevenson 2014). This means considering what information forecast is required to provide and how this information should be used. Even though the investment in forecasting usually includes some additional software or technologies, these should only be acquired after the systems outputs are defined. This includes defining needed improvements for forecasting process, estimating if new data sources or other inputs that are needed and defining what reports and outputs are expected. (Feigin 2014) Technologies rarely solve problems only by purchase and investing in forecasting technologies does not differ from this. It is vital to understand what the forecast is, for it to be used efficiently in the decision making (Stevenson 2014). Since the investment in forecasting is often expensive Feigin (2014) suggests evaluating the process and with cheap, easily accessible software. This kind of test run is good way to determine whether the investment could produce value or whether it is waste of money similar to minimum viable product in lean thinking (Euchner 2013).

3.1.1 Components of forecasts

Forecasting is important part of business. Predictions of future are needed throughout the organization. For example, demand planning requires forecast to match supply with demand and upper management requires forecasting for efficient strategic planning. (Stevenson 2014; Feigin 2014) Since the forecasting is used in various areas, techniques and outputs of the forecasts vary a lot but there are still features that generally all forecasts have in common.

Firstly, forecasting techniques assume usually that future is a reflection of past and therefore historical data is often used as the basis for the forecast. Secondly forecasts will generally differ from the reality and this feature should be accounted and in the forecast. (Stevenson 2014) As stated by Feigin (2014) in ideal situation, consolidating forecasts of individual products should lead to same forecast as the forecast of product group. Though this is the case in ideal situation, generating multiple forecasts for individual products creates forecasting errors for each product which need to be assed. Generally grouping products under single forecast is more efficient since the forecast errors tend to cancel each other out (Stevenson 2014). Fourth component, good forecasts generally have in

common is that forecast with longer time period have decreased accuracy. Future is full of uncertain variables and longer the time horizon of the forecast is, more of these variables there are. (Stevenson 2014)

Good forecast also has identifiable characteristics. Since the objective of forecast is to be used in decision making as actionable intelligence similar to results of BI&A process presented in chapter 3.3, accuracy, timeliness and reliability are also important characteristics of good forecast. (Bose 2008; Stevenson 2014) In addition to these, good forecast also has other characteristics and fourth is that forecast should be presented in meaningful units. This refers to considering the target audience of the forecast. For example, financial forecasts should use monetary metrics whereas production forecasts should use quantities as metrics. Fifth characteristic is that forecast should be presented in writing. This ensures that all users have the same information. (Stevenson 2014)

Trustworthiness of the forecast is based on whether the user understands the methods used in forecasting (Stevenson 2014). Forecast which is simple to generate and use is therefore better than forecast which uses complex methods and therefore simplicity is often emphasized in good forecast. Lastly, good forecast should result in more benefits than costs (Stevenson 2014). Feigin (2014) notes that benefits from forecasting are often indirect and therefore benefits should be considered from various point of views to identify them.

3.1.2 Forecasting different products

Product type needs to be considered when creating forecasts. As Feigin (2014) states at least three different forecasting processes should be identified depending on the product type being evaluated. These three types are stable products, new products and transition period products. Stable products refer to products which has historical data available, is going to be on the market in the near future at least and its demand is predictable. This research shall focus on the stable product demand forecasting.

New products are problematic and need to be considered as their own prediction process when forecasting. This is due the fact that new products don’t have historical data available about and market demand is often biased. Transition period products refer to products being replaced by newer models. Transition period products refer to both, models being replaced and the newer models. Demand of both needs to be considered and therefore third process for forecasting is needed. (Feigin 2014)

In the scope of this research demand can be further divided into continuous and discrete demand. Continuous demand is constantly ongoing where discrete demand occurs more scattered. Continuous demand is a result of customer’s stable requirement for a standardized product where discrete demand occurs for example in projects where specific product is needed. Historical data can be used as source of forecasting for both demand types but

in the case of discrete demand, other information sources such as sales personnel expertise should be used to enhance the forecast accuracy (Stevenson 2014).

3.1.3 Creating forecast

Armstrong (2001) introduced the six major steps for the organization when creating their forecast. The process is considered from very vast point of view and is meant to be defined specifically for the use. Armstrong's process for forecasting is presented in the Figure 1.

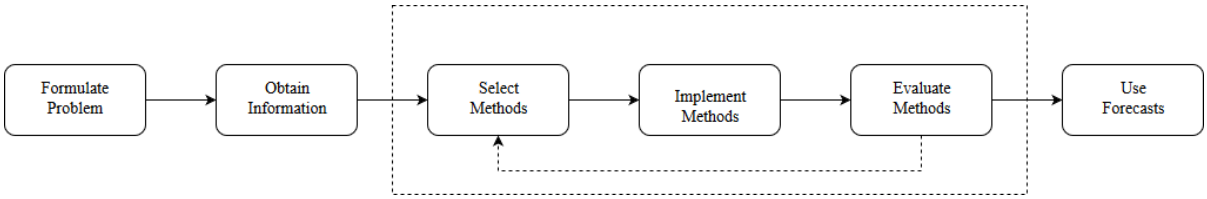

Figure 1. *Stages of forecasting (Armstrong 2001)*

Stages of forecasting presented by Armstrong (Armstrong 2001) start with a problem formulation. This step requires to understand the problem being forecasted. Understanding the problem is crucial and if this step is over looked, outcome of the process might not answer to the problem that it was intended to provide answer to. In the second step the forecaster needs to obtain information relevant for the problem. If there is not enough relevant information available, this will affect the outcome of the process. On the other hand, high quantity of information might hinder the process if unnecessary or incorrect information affects the results (Laihonen et al. 2013).

Selecting, Implementing and Evaluating Methods are presented as iterative process within the whole process (Armstrong 2001). After the forecasts are considered as viable organization will in the final step use these forecasts until sequential forecasts are published. Since forecasting reports are often requested on ad-hoc basis as descriptive analytics time spent for the process needs to be considered (Kaisler et al. 2013). Often the report is needed swiftly and therefore too much time should not be spent on iterating the most accurate forecast. Feigin (2014) notes that perfect forecast can't be achieved and time spent creating the forecast should be considered versus the estimated value gain from the forecast.

Feigin (2014) presents more detailed forecasting process. When Armstrong's process is considered as an overview of all forecasting processes Feigin's version is created especially for forecasting stable products. Process is divided into five main steps but as Feigin recognizes the importance of determining whether historical data can be used as the basis essential part it is identified as the second step in the Figure 2. Starting point should be

considered to be determining the purpose of the forecast. Business need behind the forecasting can be argued to be the starting point in the forecasting process. Understanding the business need requires understanding what is expected from the forecast as well as what is not expected from the forecast. This requires definition of the models to be used in the forecasting as well, determining the data sources and the time horizon of the forecast. (Stevenson 2014)

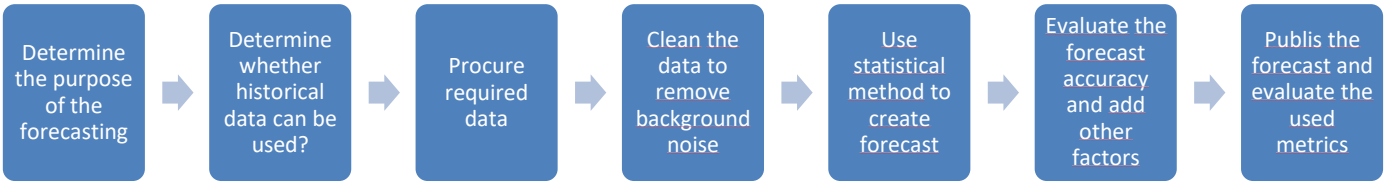

Figure 2. *Forecasting process for stable products (Feigin 2014)*

Feigin's process aims to produce unbiased forecast and determining what data sources should and can be used are important basis for the process. Data procurement should be executed based on the expected result since for example forecast based on historical data on orders will produce different result as historical data from realized shipments (Feigin 2014). Armstrong (2001) notes that shorter time horizon requires less data for accurate forecast. Therefore, forecaster should not overlook evaluating relevancy of the data sources and whether it is probable that future will even reflect history.

Third step, cleaning the data from the background noise can be seen as the most important and repeatedly overlooked step of the whole process (Feigin 2014). Often, data is inaccurate and contains misleading abnormalities which will lead to error in the forecast (Armstrong 2001). These abnormalities are result of internal or external causes which affect the demand of the product. For example, affects from marketing campaign which has caused spike in the sales create biased forecast on organization's historical data, but they might also create pits in the competitor's historical data. Armstrong (2001) suggests mathematical methods to be used in identifying the abnormalities. These methods could be for example setting high and low baseline for the demand forecast of the product or checking the data points against means or standard deviation.

When Armstrong's (2001) process focuses more to the finding appropriate statistical method for the process, Feigin (2014) sees the use of the statistical method as smaller part of the process. Effect of trend and seasonality should be first removed from the data to create the baseline value for the demand. Afterwards the effects of both are added to the baseline value. Influence of trends are generally added as separate values to the baseline demand, but seasonality is usually presented as index which affects the baseline value.

Seasonality varies depending on the product. For example, clothing industry demand varies vastly depending on the season when stock market has much less seasonality (Armstrong 2001). Based on the historical data organization should identify number of seasons in their business and create separate seasonal indices for each (Feigin 2014). Impact of the seasonality should be considered before blindly adding it to the baseline demand. Identifying seasonality is often challenging and especially when the market situation is unstable it can be the cause of biased forecast, in which case the effect of the seasonality should be reduced (Armstrong 2001).

Trends are easy to spot when they are happening, but it becomes harder when trying to determine when they are going to change. Most common mistake when calculating the effect of trends is to over or underestimate the affect and therefore, when creating forecasts, there should be versions with and without the effect of the trend (Feigin 2014). When trend is miscalculated the forecast without trend will perform better and forecaster has a continuance forecast ready. When situation is ambiguous or the time horizon of the forecast is long, effect of trends should be treated cautiously (Armstrong 2001).

Evaluating forecast accuracy and other factors is important step before publishing and using the forecast. Wrong or misleading information in forecasts can lead to misguided decisions and financial loses. At this point statistical forecast should be evaluated, and the effect of additional factors should be included if necessary (Feigin 2014). Additional factors refer to facts, which would suggest that the forecast is wrong and should be corrected. Armstrong (2001) notes in many occasions that domain knowledge and expertise of the forecaster should be used throughout the process. Evaluating the process emphasizes this. Evaluating the forecast should not be done if there is no reason to suspect that the forecast needs to be adjusted (Feigin 2014).

Forecasting evaluation can be done on detailed level or the process can be done minimal effort. Simply this could be done by concentrating the evaluation to products which have had bad forecast realizations in the past and to products which have significantly over or under performed in their previous forecasts. (Feigin 2014) Since relevant data sources should be utilized when creating forecasts, in the demand forecast customers can provide valuable information on the forecast accuracy (Armstrong 2001). When evaluating the forecast more precisely, customers view should be considered. Simple way to initialize information sharing can be just to present organization's forecast of customers demand to their customer and ask whether they think it will hold. When the cooperation regarding information sharing deepens, modern business intelligence tools offer interactive information sharing which can be easily utilized to share forecasting reports and allowing customers to correct them based on their own view. (Feigin 2014)

The final step of Feigin's (2014) forecasting process is evaluating the forecasting process and publishing the forecast. Process evaluation should include analyzing the metrics used in the forecast. Without understanding the process and metrics, forecaster can't be certain

that they have actually answered to the original problem (Armstrong 2001). Generally organization's don't spend enough time on interpreting the forecast and without understanding the phenomena, the process can't be developed (Feigin 2014). Evaluating the process also assures that at least the biases of the forecaster have been inspected which increases the reliability of the forecast

When the forecast is deemed ready, it should be shared to all parties who will need the information, so the information is available for decision making. On the publication, the forecast should be presented in easily understandable, meaningful form keeping in mind that it will be the basis of decision making. Unbiased forecast aims to deliver objective view of the demand and not reflect forecaster's own vision of the situation (Feigin 2014). Precise numbers should also be avoided since they might generate overconfidence in the forecast's accuracy (Armstrong 2001).

3.1.4 Forecasting methods

Armstrong (2001) divides forecasting methods into two main approaches based on whether judgmental or statistical knowledge source is used. Judgmental knowledge source, more commonly known as qualitative approach to forecasting uses judgement and opinions to produce forecast (Stevenson 2014). As noted by Feigin (2014) qualitative approach should be used especially when organization considers that future will not be a representation of historical data. This is often the situation in case of new products or if up to date historical data simply unavailable (Stevenson 2014).

Qualitative approaches can be further divided based on whether the analyst is forecasting their own future or creating forecast of someone else's future (Armstrong 2001). Knowledge sources used in these cases can be internal or external from the organization's point of view. Generally qualitative approach is applied by collecting knowledge from upper managements view of the future, sales personnel views of the future sales or collecting information straight from the customer. Other data sources can also be used such as other staff of the organization or outside consultants for example. (Stevenson 2014)

Elements of forecasting methods and their relationships are potraited in the Figure 3. Judgemental and statistical approaches are seen in the first level as far away from each other but in case of high-level forecasting they are supplementary to each other and more integration beteen judgmental and statistical approaches is needed. (Armstrong 2001) As noted by Feigin (2014), it is advantageous to use two methods to validate the forecast. In practice this could be done for example in case where forecast is needed switfly by creating judgemental forecast first and validating this with statistical approach (Stevenson 2014).

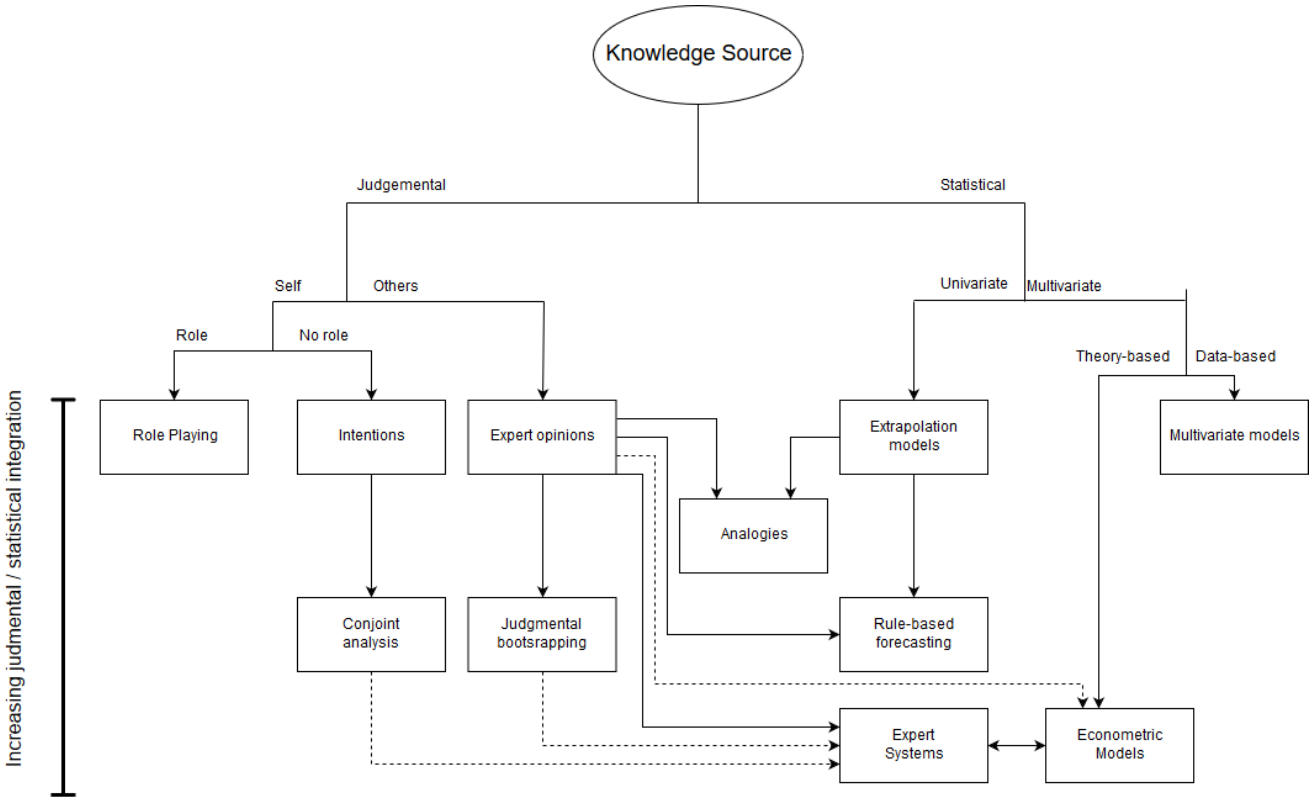

Figure 3. *Elements of forecasting methods and relationships (Armstrong 2001)*

Statistical, or as it is often referred, quantitative approach is the second commonly utilized forecasting method. Quantitative approach utilizes analyzing of raw data which is processed and analyzed to produce forecast. (Stevenson 2014) Unbiased forecast often utilize quantitative knowledge sources since raw data makes it easier to disregard forecaster's own biases (Feigin 2014). Continuous demand reflects historical data and can therefore be forecasted using statistical methods. Forecasting discrete demand requires usage of other sources in addition to historical data and therefore judgmental forecasting methods should be used. (Armstrong 2001)

Selecting the appropriate forecasting method requires considering numerous factors. Factors to be considered are for example available data, pattern of the data, forecasting time horizon, time required to execute the forecast and the skillset of the forecaster (Stevenson 2014). All of these factors are important but as a generalization, it can be said that the most important factors to be considered are accuracy and money (Feigin 2014). They are two opposite sides of the forecasting starting point: higher accuracy generally means more costs. Table 2 summarizes the elements and suitability of forecasting methods.

Forecasting method	Amount of historical data	Data Pattern	Forecast time horizon	Time required for prepartion	Skillset of the forecastor
Naive forecasting	1-2 observations	Stationary	Short	Short	Basic understanding of forecasting

Moving average	2-30 observations	Stationary	Short	Short	Basic understanding of forecasting
Simple exponential smoothing	5-10 observations	Stationary	Short	Short	Basic understanding of forecasting
Trend-adjusted exponential smoothing	10-15 observations	Trend	Short to medium	Short	Advanced understanding of forecasting
Trend models	10-20 observations, for seasonality at least 5 per season	Trend	Short to medium	Short	Advanced understanding of forecasting
Seasonal	At least 2 observations of the season being forecasted	Cyclical and seasonal patterns	Short to medium	Short to moderate	Basic understanding of forecasting
Causal regression models	10 observations per independent variable	Complex patterns	Short, medium or long	Long development time, short to implement	Expert understanding of forecasting and business variables

Table 2. *Selecting the appropriate forecasting method (Stevenson 2014) p.108*

As can be seen from Table 2 selection of forecasting method depends highly on the situation. If only stationary data is available this creates a situation where only simple forecasting methods can be utilized. Secondly expertise of the forecaster might limit methods. As noted by Stevenson (2014) p76., it is important that the forecaster understands used techniques. Especially understanding the circumstances and the limitations of the forecasting methods are often overlooked in this case.

Generally most used forecasting methods are naïve forecasting, moving average and various methods of exponential smoothing when forecasting is done with quantitative methods (Stevenson 2014) pp. 82-87. These methods are presented introduced below.

NAÏVE FORECASTING

Naïve forecasting is the most commonly used method for quantitative forecasting. Naïve method can be used when forecasting stable series either with or without trend and seasonality (Stevenson 2014). In naïve approach, base value is simply the value in the measured in the previous point. If seasonality is included the forecasted value equals to value of last same season. Trend is calculated only based on the latest two values, where the latest value produces the baseline for the forecast and the effect of the trend equals the difference between the last two recorded values which is then combined with the base value.

Naïve approach in its simplicity is widely used as forecasting method. It has many qualities which good forecast should have but the two main points are that it is easy to understand & execute, and it is cost-effective since the creation of naïve forecast can be done with minimum effort and resources. (Stevenson 2014) Accuracy is the biggest weakness of naïve forecast and as stated by Feigin (2014) organization should always consider how much they are willing to invest to improve forecast accuracy and what is expected and required from the forecast in order to determine if current method used is acceptable for them. This is also true when utilizing naïve approach. If it produces acceptable results, it can be enough for the organization.

MOVING AVERAGE

Data often contains some abnormalities, for example as a result of marketing campaign (Feigin 2014). These cause often variation in the data distribution and since identifying the reason and removing its affect from the data is time-consuming task, moving average allows smoothening of the forecast (Pihlajamaa 2015). Moving average is the average value of chosen number of previous datapoints.

Choosing the number of data-points affects the sensitivity of forecast. When more data points are used, the forecast does not react easily to changes trends and on the other hand, when only few data points are used, effect of trends is easily overestimated (Stevenson 2014). This creates the disadvantageous of the moving average: all data points are valued equally and when using only few data points, reactiveness emphasizes in the forecast in the expense of losing information from further datapoints and vice versa in case of using more data points. Calculating moving averages is simple and the forecast is still easy to understand but it takes a bit more effort from the forecaster compared to naïve method. (Stevenson 2014)

EXPONENTIAL SMOOTHING

Exponential smoothing is a forecasting method that emphasizes the importance of newer data points without disregarding older points. Forecasted value is calculated based on previous forecasted value to which is added difference between realized value and the forecasted value multiplied by the smoothing constant. Smoothing constant adjusts the forecast based on the forecast error. It is a value between 0 and 1. Smoothing constant zero adjusts to the forecast error slowly and if the constant is 1 faster the adjustment is made. Smoothing allows adjusting the forecast depending on the situation and because of this and the fairly simple mathematics, exponential smoothing is also practical and often used forecasting method. (Stevenson 2014)

If trends are needed to include in exponential smoothing their value needs to be calculated separately. Exponential smoothing is also widely used forecasting method since it is fairly simple, and many business intelligence and forecasting tools enables it's use with minimum effort. (Feigin 2014; Stevenson 2014) There are also more sophisticated version of

exponential smoothing forecasting such as Holt-Winters method which includes seasonality and trends to the forecast (Pihlajamaa 2015). In Holt-Winters, to include seasonality it is suggested to have at least three years of data available. This is often referred as minimum requirement since even then forecaster only has three data points to which base the seasonality which is not statistically that significant. (Feigin 2014)

3.2 Business intelligence and analytics

3.2.1 Data, information, knowledge and wisdom

Business Intelligence and Analytics aim to process raw data into information which can be used in decision making providing actual benefits for organization (Duan & Da Xu 2012). Knowledge management addresses issues regarding data and information refinement and on the other hand issues regarding creating new data and information as well as managing data and information (Laihonen et al. 2013).

Defining business intelligence (BI) is difficult since there is countless definition available depending on the focus of business intelligence. In the scope of this research, BI is defined in the scope of its purpose which is to actively seek information discoveries from data, to help decision making and find insight to create value for business (Liebowitz 2006; Duan & Da Xu 2012; Laihonen et al. 2013). Diversity of definition is partly because definition of data and information is also ambiguous. Generally used way to understand them is to use data, information and knowledge as levels which represent different levels of understanding the data. These levels can be complemented with intelligence, wisdom and truth which represent higher levels of understanding. (Thierauf 2001; Laihonen et al. 2013) Figure 4, represents the levels of understanding and the relationships in between the levels.

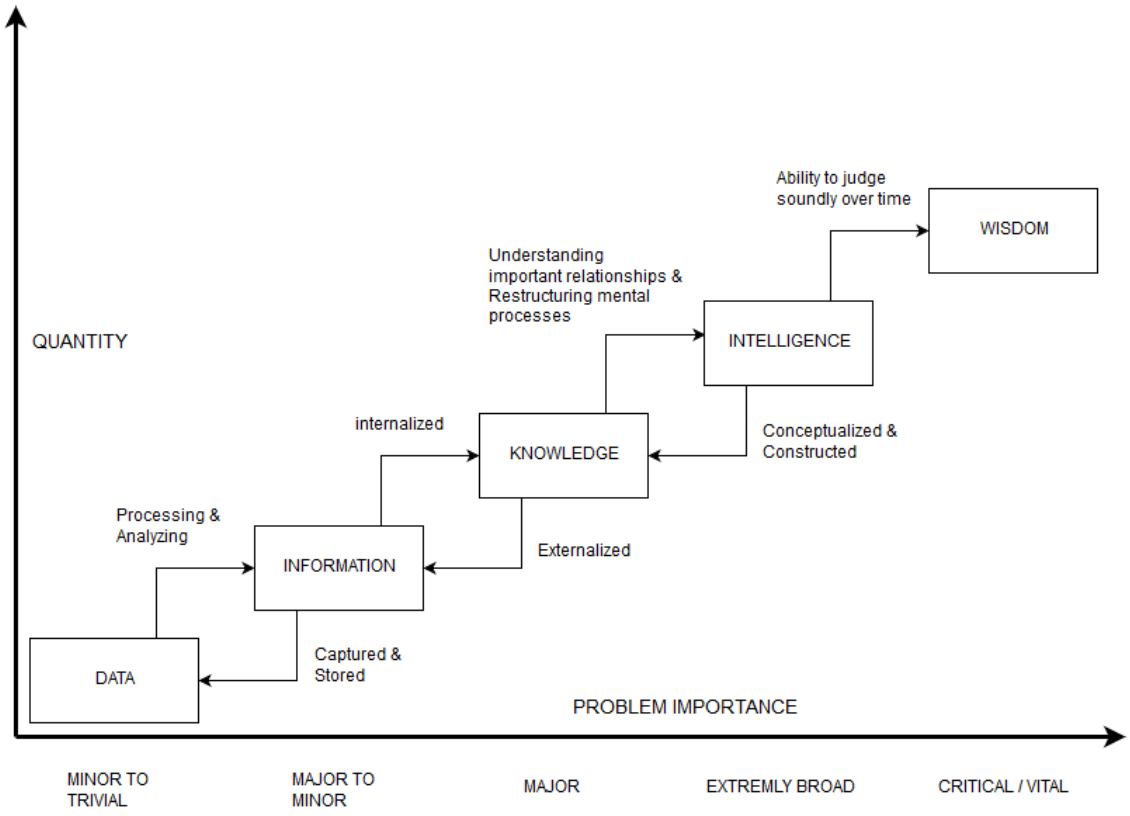

Figure 4. *Levels of understanding and their relationships, modified from(Thierauf 2001; Liebowitz 2006; Liew 2013; Laihonen et al. 2013)*

On the lowest level in the Figure 4 is data. Data is collection of unstructured facts or symbols which by themselves produce little value and can mainly be used to answer only minor problems (Thierauf 2001; Duan & Da Xu 2012). As noted by Duan & Da Xu (2012) one of the main purposes of data is to act as storage for actions. Typically, data processor should aim to move beyond data-level to produce value (Thierauf 2001). This fact should also be comprehended by organization's management when considering investing in data since the investment in data produces value only after data is processed into actionable level (Loshin 2012).

Data is refined into information by processing and analyzing the data (Duan & Da Xu 2012). Simply put data can be seen as information as soon as it is patterned somehow (Liebowitz 2006). Information is itself valuable and it can be used to solve mediocre problems by using information as the basis for the decisions. The problem is that information is still low-level understanding and there are often vast quantities of information available related to the subject which won't simplify the decision making. (Thierauf 2001) Therefore moving towards deeper understanding should be considered for broader perspective and to avoid biases. It should be noted that information can be processed back to data level when it is captured and stored (Duan & Da Xu 2012).

On the next level is knowledge, which is further develop and internalized intelligence (Liebowitz 2006). Knowledge is seen as expertise gained from experience which is combined with additional information on the subject and it can be used to tackle major problems. Knowledge itself can't be transferred easily and it is assimilated within the expert. It is formed in the mind of the information processor. For example, information about organization's finances can be seen as knowledge in the hands of organization's financial department which can combine it with their own experience on the financial sector. (Thierauf 2001) As simplified by Duan & Da Xu (2012) information is converted to knowledge by internalizing information and on the other hand knowledge can be converted back to information by externalizing.

Higher levels of understanding: intelligence and wisdom are often considered as single level which is wisdom (Duan & Da Xu 2012). Because intelligence contains characteristics and relationships with knowledge and wisdom it is advantageous to be considered as its own level (Thierauf 2001; Duan & Da Xu 2012). Intelligence is understanding various relationships between data, information and knowledge targeting to give insight or answer desired objective or question. Intelligence on the subject should provide enough understanding for the decision making for the most cases and therefore higher levels of understanding is rarely required in the organization (Thierauf 2001). It should be noted that intelligence itself can't be converted back to knowledge level (Duan & Da Xu 2012).

Wisdom has multiple definitions but generally it is considered to be the highest level of understanding (Duan & Da Xu 2012). Though some research indicates that truth could be seen as highest level after wisdom (Thierauf 2001; Laihonen et al. 2013), in the scope of this research wisdom is considered to be the highest level on understanding. It can be said that judging the situation and understanding what data, information, knowledge and intelligence is needed can be considered as wisdom (Thierauf 2001).

3.2.2 Business intelligence

Business Intelligence (BI) is a set of tools, technologies and processes which all aim to provide information to endow decision-making (Davenport 2006). Watson & Wixom (2007) define two main activities of BI: get data in and get data out. The process in between, includes various other steps depending on the information needs but in the end purpose of BI is to produce correct, relevant and well-timed information to support decision-making in organization (Davenport 2006; Watson & Wixom 2007; Chen et al. 2012; Laihonen et al. 2013).

Literature often shows BI as ambiguous term due its many dimensions. One way to manage the definition of the term is to consider its purposes. In the case of BI Gilad & Gilad (1985) have defined following purposes for it:

1. Obtain data, information and knowledge from various sources

2. Store data, information and knowledge
3. Evaluate data, information and knowledge
4. Analyze data information and knowledge
5. Disseminate to relevant parties

As stated by Vuori (2011) these purposes of BI are not necessarily connected to each other and they can be executed separately. For example, organization can collect data without knowing whether it leads to actions or on individual level, employee might exercise many BI operations in their job but don't share this information to others.

Obtaining data from various sources proves often to be the most challenging feature in BI. As a result, it takes roughly 80% of the time in BI and is often identified as reason for unexpected expenses in BI projects. (Watson & Wixom 2007) Even though getting data in, consumes much time it is important for effective BI. As overwhelming amount of data is available (Chen et al. 2012), it is important to relevant high-quality information from the data. Excessive amount of data is unnecessary and can even be harmful if relevant information can't be separated from it (Laihonen et al. 2013).

Results of effective BI are not only in the dashboards shown to management or other decision-makers but also in the collected information which can be utilized also in the future (Laihonen et al. 2013). This defines the job description of employees working in the field of BI: collecting, analyzing and storing data for future use is as important as producing reports or dashboards from that data. Efficiently stored data can for example be used to create dynamically dashboards in which the data is only minutes old (Watson & Wixom 2007).

As stated by Thierauf (2001) BI systems have generally gathered information from organizations internal functions such as marketing, sales and manufacturing. Processing external information has often been seen as part of competitive intelligence (CI). Competitive intelligence focuses to external issues such as competitor analysis (Liebowitz 2006). In this research, BI is defined through its purpose: providing actionable intelligence for decision making to create value (Bose 2008). Therefore, BI is seen as the top-level term for intelligence analytics with both internal and external factors included and by extension CI is sub-category under BI.

3.2.3 Evolution of the Business Intelligence & Analytics

BI is often difficult to define from a single point of view. In this thesis BI is partly processes and partly technologies. Use of these requires various analytical skills as well and therefore, more comprehensive term Business Intelligence and Analytics (BI&A) is used to define this thesis' scope. BI&A is collection of techniques, practices, technologies and systems that all analyze the market and aim to help in timely decision making in business (Chen et al. 2012).

Business Intelligence and Analytics (BI&A) has evolved majorly when the possibilities of data and analysis have been realized. Database management can be seen as roots for BI&A (Chen et al. 2012). Data sources have evolved from organization's internal structured databases to various external sources. Inputting data to internal databases is still important activity and analysis of sales data can predict future sales but collecting data for example through cookies in ecommerce can offer complementary information about customers preferences (Chen et al. 2012).

Chen et al. (2012) identify three stages of BI&A. These three stages are seen to represent evolution of BI&A and by understanding differences organization can evaluate their own BI&A maturity. Data management and warehousing is seen as foundation for BI&A 1.0. (Chen et al. 2012). Typically, Business Intelligences systems can produce various dashboards and other data representations from structured data sources. Data mining and statistical analysis are done enhance the data source and find needed values. Leading vendors of commercial BI Systems such as Microsoft and Tableau Software include these technologies in their products (Sallam et al. 2011).

The development of Internet has provided new opportunities for data collection and analytical technologies. Piles of data from customers, competitors, or market can be collected globally using web intelligence and analytics which has opened the research for BI&A 2.0 (Chen et al. 2012). Problematically this data is often unstructured and requires more comprehensive understanding of analytical technologies from the data processor. Social media for example has offered ways to develop traditional Business to Consumer (B2C) market towards interaction and discussion between organization and customer. To understand this kind of change in business, research is needed and unlike BI&A 1.0 various mature IT and BI tools need integrations to enable future BI&A 2.0 tools to work efficiently with existing systems. (Chen et al. 2012)

Even though many organizations are yet to integrate BI&A 2.0 systems to their business research has already turned to BI&A 3.0. Where BI&A 2.0 data sources are based on user's activity in Internet and for example automatic content analytics in BI&A 3.0 data is collected through personal sensors such as smartphones (Chen et al. 2012). Internet of Things (IoT) has opened new possibilities for business and many organizations have their own IoT product in development. Context, omnipresence and optimization which are the three separable attributes in IoT offer vast collection of data for organizations to use. (Witkowski 2017). Successful commercial applications for this data is still in a few and efficient BI&A 3.0 systems are not in general use in organizations (Chen et al. 2012).

3.2.4 Levels of analytics

Need for analytics varies greatly in organizations and as a term analytics is used to cover vast amount of decision support processes and optimization models (Kaisler et al. 2013). Thus, analytics should always be done based on business needs to produce most efficient

result for the need. Generally, analytics are divided into three levels: descriptive, predictive and prescriptive analytics based on the purpose of the analytics (Idoine & Herschel 2016; Wang, G. et al. 2016; Wang, C. et al. 2018). The levels of analytics should be seen cumulatively, and organization needs to possess basic analytics before aiming for more difficult analytical processes (Idoine & Herschel 2016). Additionally, diagnostic analytics is often identified as a level of analytics between descriptive and predictive analytics (Kaisler et al. 2013). These four levels are defined as analytical levels for this thesis. Figure 5 represents the business value each level can offer and how much human action is needed before decision making and action.

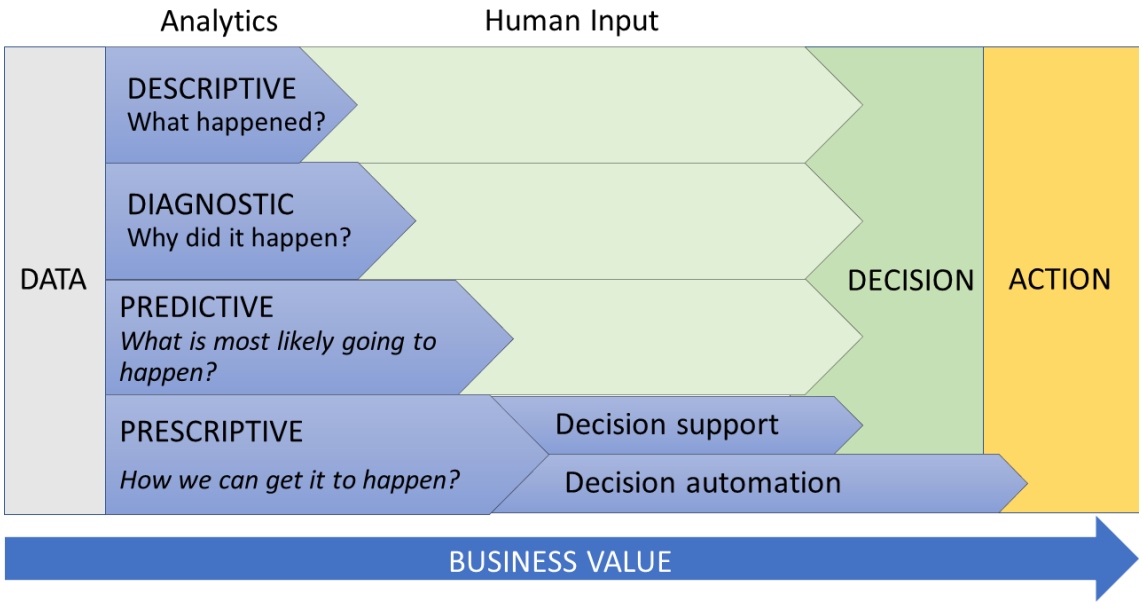

Figure 5. *Value of analytics to business (Idoine & Herschel 2016)*

Descriptive analytics is the most basic form of analytics. Techniques of descriptive analytics look in to history and aim to answer the question "What happened?" This is done with techniques which aim to understand existing data (Kaisler et al. 2013). Dashboards and other visualizations which are done for reporting is typical example of the descriptive analytics. Descriptive analytics is done on ad-hoc basis for example when management needs it (Wang, C. et al. 2018) or descriptive analytics anal can also be done on standardized times such as organization's financial releases (Wang, G. et al. 2016). Most organizations do these kinds of basic analytics to communicate their financials for example.

Diagnostic analytics is the second stage of analytical techniques. In diagnostic analytics the techniques are more complex and understanding the phenomena or question under examination is more important than understanding the data itself. Diagnostic analytics aim to answer to the question "Why did it happen?" (Kaisler et al. 2013) Like descriptive analytics, diagnostic analytics also needs much human input before the data can actually be used to create value for business (Idoine & Herschel 2016). Even though results of diagnostic analytics are often presented similarly as results of descriptive analytics the

techniques used to create these are more complex. Diagnostic analytic techniques use different data sources used to find correlations in the data and the result is not just a visualization of historical dataset (Kaisler et al. 2013).

Predictive analytics seek the answer to the question "What is going to happen?" or more precisely "What is most likely going to happen?" since all predictions of the future are based on the information available and no prediction can promise that something is going to happen (Idoine & Herschel 2016). Predictive analytics includes techniques used in both levels mentioned above, but it also utilizes mathematical methods such as statistical analysis, data mining and machine learning to create forecasts of the future (Wang, G. et al. 2016; Wang, C. et al. 2018). Depending on the organization's needs predictive analytics can be seen also as a way to process data to enhance business (Idoine & Herschel 2016).

Prescriptive analytics is a set of analytical techniques that aim give advice to decision makers on how to reach business targets. Prescriptive analytics aims to answer to the question "How we can get something to happen?" (Idoine & Herschel 2016) Understanding dependencies and correlations between various objects is important in prescriptive analytics. Per Wang et. al (2016) one of the key characteristics in prescriptive analytics is assessing alternative decisions. This can be approached with techniques such as machine learning and complex event processing (Idoine & Herschel 2016). Running these kinds of complex mathematical processes requires computing power and understanding of the mathematics behind them. Organization need to be committed to high level analytics to get value of using prescriptive analytics. Otherwise the implementation of high-level analytic tools might lead to considerable financial loses (Wang, G. et al. 2016).

When aiming for specific level in analytics it is important to understand organization's own maturity in analytics. If the maturity of analytics does not support prescriptive analytics, implementing complex systems such as machine learning platforms most likely will only lead to unnecessary expenses which won't support organization or the business at all. Analytics are in the end only a set of tools which aim to create value for business and this requires efficient use of the tools. (Idoine & Herschel 2016)

3.3 Business Intelligence & Analytics process

Process used in business intelligence and analytics has been introduced in various literature about generic intelligence and more specifically in literature about business or competitive intelligence (Thierauf 2001; Bose 2008; Vuori 2011; Fleisher & Bensoussan 2015). Process is typically presented as continuous cycle from where intelligence creates new business needs which then restart the whole cycle. Table 3 presents phases of the cycle and tasks associated with them.

Phase of the process	Typical actions during the phase
Defining information needs	- Determine information needs: what, when and why (internal & external) - Determine who needs what information in the process: crucial information should be restricted based on actual need.
Collect data and initial processing	- Collecting data from required sources (internal & external) - Evaluating the data quality, practicality and reduction of unnecessary data - Classification of the data (identifying patterns, dependencies and anomalies)
Analyze the data	- Using different methods to analyze the data, information and knowledge collected to produce intelligence - Storing collected data and information for further use
Disseminate gained intelligence	- Sharing the information to relevant parties - Deliver finished report or intelligence to decision makers
Evaluate and utilize	- Decision making and action based on intelligence - Collect feedback of produced intelligence (was it used, why how etc.) - Determine whether new intelligence needs resulted due the process

Table 3. *Typical phases of BI process, concluded based on (Thierauf 2001; Bose 2008; Vuori 2011)*

Defining information needs phase is often identified as the starting point for BI&A process. Getting started with the process can be one of the most challenging steps of the whole process and often organization aren't able to determine what their needs are. View from outside source can be helpful when organizations begins executing BI&A process. Fleisher & Bensoussan (2015) endorse using outside help since they present the generic intelligence cycle from consultants point of view. All the relevant parties should be included in definition process but especially the decision makers involvement is crucial at this point so the intelligence needs that are required for the decision making are defined and the process will produce applicable intelligence (Bose 2008).

Bose (2008) suggest that defining the information needs should be done by answering questions what, why and when. Each question should be considered from the viewpoint of the analyst who will execute the process and from the point of view of decision makers. Vuori (2011) includes also question, what kind of decisions are made based on the information, in the first step to more precisely understand the needs behind the process. Definition step should include iterations, so all sides know the starting point for the process.

As stated by Vuori (2011) the collecting information consists of to main phases. Firstly, data needs to be acquired from various sources and secondly acquired data should be evaluated regarding quality, reliability and usefulness. BI&A process can include data from both internal and external sources (Thierauf 2001). The actual need behind the process should be considered when acquiring data so relevant data sources are used. In this

case relevant data sources refer to use enough data sources that intelligence can be produced but not to excessive amount so that high volume of data would hinder the process (Laihonen et al. 2013).

When acquiring information organization should always consider their own employees as a source. Employees generally have a lot of domain knowledge of the organization and their business (Feigin 2014). Vuori (2011) also emphasizes the use of human sources since they possess crucial information regarding business intelligence. Challenge of using human sources comes with cognitive biases. Biases affect decision making and judgment of the informant which will lead to emphasizing specific aspects. Since the viewpoint of most informants are biased anyways, the solution is to acknowledge the biases, analyze their affect and collect data from various point of views. (Soll et al. 2015)

Analyzing the data is the step where intelligence is produced. Since the objective of the BI&A process is to produce intelligence which would create value for business, data analyzation should lead to recommendation for the decision makers (Bose 2008; Bensoussan & Fleisher 2012). Creating intelligence requires analyzing and internalizing data and information. Understanding entities and relationships between knowledge sources is required for the process. (Liew 2013)

Analysis could be executed in various technical ways such as modelling and forecasting depending on the collected data. Quantitative data can be analyzed by utilizing simple statistics such as means and variances or more complex methods such data mining or machine learning (Vuori 2011). Qualitative approach can be used to enhance the understanding of external factors. Generally well-known methods such as SWOT, BCG Growth Matrix or Scenario Analysis are good starting point for this (Bensoussan & Fleisher 2012; Fleisher & Bensoussan 2015).

Analyzing the data or other inputs is difficult process where problem is broken into its key components and these key components are aimed to be understood individually (Bensoussan & Fleisher 2012). Complex issues require to be analyzed at smaller parts to understand more than just the most obvious answer (Bose 2008). Analyzing requires understanding the data and evaluating its reliability and accuracy (Vuori 2011; Bensoussan & Fleisher 2012). As stated by Bensoussan & Fleisher (2012) sufficient data collection is a good starting point for analysis. If this is not the case analyst should consider collecting additional data. Analyzer should consider usefulness of the data, information and knowledge (Vuori 2011). Without considering usefulness the gained intelligence might not resolve the original problem and therefore, when evaluating usefulness original problem should be considered.

Disseminating the information to decision makers and other relevant parties is the actual benefit and result of the process. Gained intelligence can be used in decision making or in further analysis. (Bose 2008) Dissemination often takes form of dashboards or reports

which should be constructed keeping in mind the target audience. Results should be communicated and made available for all parties who might benefit from it (Vuori 2011).

Last step in the process is to evaluate and utilize the information. When intelligence produced is provided timely and accurately to the decision makers it provides deeper understanding on the subject which is then basis for decision making and action (Vuori 2011). Second part of the final step is to collect feedback from the parties involved in the BI&A process. This assures that the process satisfied the original intelligence requirements and acts possibly restarts the process if that is not the case (Fleisher & Bensoussan 2015). It is also important to understand whether the intelligence produced was used and how (Bose 2008). This will allow modifying the process for future use and possibly identifying new information needs so the cycle can start over again.

Much of the literature regarding BI&A process note that it is important to understand the business neeed to produce actionable intelligence which can be used to create business value (Bose 2008; Vuori 2011; Bensoussan & Fleisher 2012). The business value is created through analysis process and for effective analysis, organization needs to be committed to it. Commitment is required especially from management whose support is necessary to advance the role of BI&A. From employees, this requires understanding the fundamentals of analytics as well as the business needs. (Díaz et al. 2018) Therefore, business need can be seen as part of the BI&A process and as the trigger for the whole process.

In the figure 6, BI&A process is presented from the viewpoint of this research. Process is constructed based on the general intelligence cycle e.g. presented by Fleisher & Bensoussan (2015) and competitive intelligence cycle e.g. presented by Bose (2008). Process also includes the purpose of analytics: to make better decisions (Díaz et al. 2018). Therefore, the business need is seen as the trigger of the process which should be kept in mind throughout the whole BI&A process.

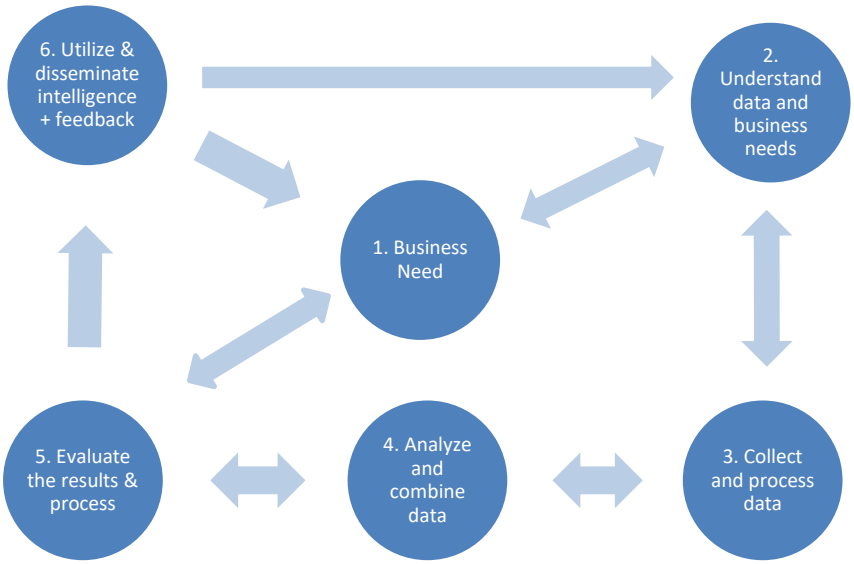

Figure 6. *Business intelligence & analytics process. Modified based on (Bose 2008; Fleisher & Bensoussan 2015; Díaz et al. 2018)*

First step of the BI&A process is the business need for business intelligence and analytics. In this step problems or expected benefits from executing BI&A should be evaluated determine whether triggering the whole BI&A process is going to be worthwhile. Second step, understanding the data and business need, refers to defining the data requirements based on the business need. Defining the data requirements is important part of the process so appropriate amount applicable data can be collected and processed in the following steps.

Third and fourth steps of the BI&A process are similar to collecting and initially processing the data & analyzing the data presented in the Table 3. BI&A model presented in figure 6 additionally emphasizes the relationships between the steps. Analyst should consider moving back in the process as well as forwards if some actions were missed in the step before. Evaluating the intelligence is identified as its own step even though continuous analyzing of the usefulness, reliability and quality of the data, information and knowledge should be evaluated throughout the process (Vuori 2011). Importance of evaluating the results and if necessary doing additional analyze should be done before the intelligence is distributed decision makers. the original business need behind the process should be kept in mind and therefore the relationship with the fifth and first step is emphasized in the figure 6. (Díaz et al. 2018) The process evaluation should also be executed in the fifth step from the analyst's point of view.

The sixth step completes BI&A process with sharing the information with relevant parties to be used for actionable decision making. It's also vital to remember to collect feedback from the process. Based on the feedback, process can either restart from the second step

if the intelligence produced is insufficient or if additional intelligence needs have risen the process goes back to the first step.

3.4 Proposed analytic process framework

3.4.1 Requirements for the operators

Understanding the business need is vital part of the BI&A process. In the forecasting context all parties need to understand what is expected from the forecast. This requires that business need is clearly defined and communicated with all parties who take part in the forecasting process. As stated by Davenport & Patil (2012) data scientist can create and identify meaningful structures in vast data masses to serve business need. As stated in the chapter 3.3 organization might struggle with understanding the actual need behind the BI&A. Therefore, the forecaster is required to have a strong connection to organization's business to identify the actual needs rather than conducting forecasting project just based on the assignment.

Chapter 3.1.3 noted that it is important to understand the purpose of the forecast. What is expected from it and what it is used for. Often forecasts are made to support decision making but if management expects decision making based on the forecast more effort is needed to create the forecast (Feigin 2014). Mutual discussion between forecaster and business representatives should be carried out through BI&A process when creating forecasts. In this forecaster's role can be seen as data scientist who aims to build things from the data to produce value for business. They need to be able to communicate the capabilities of data to the business (Davenport & Patil 2012).

Forecasters understanding of different methods sets the conditions for the hole forecasting process as noted in the chapter 3.1.4. As a basic rule, forecasting method which is understood by the forecaster will produce more accurate results than a misunderstood complex method. This affects BI&A process especially in the beginning when definitions and scope of the process is defined. If the forecaster doesn't understand selected methods, analyzing and combining the data is likely to take much longer time than anticipated when more understanding about the methods is needed to be learned. Understanding of various methods both quantitative and qualitative is therefore one of the most important skills of analyst executing BI&A process and forecasting (Schoenherr & Speier-Pero 2015).

Time horizon of the forecast is also noted as variable which affects accuracy and therefore expenses of the forecast (see chapter 3.1.4). When time horizon increases, additional data sources are often needed to enhance the accuracy of the forecast. Judgmental sources offer often untapped and these sources should be utilized as noted in the chapter 3.3. Management and other employees can offer judgmental data sources for the forecasting and therefore lead to enhanced forecast accuracy. Especially management input is often

needed when time horizon of the forecast increases but this will also lead to increased costs of the forecast (Stevenson 2014) p. 105.

Important skill of data analyst is to understand business and its financials (Schoenherr & Speier-Pero 2015). As noted in the chapter 3.1.1 benefits should be greater than costs in a good forecast. As there are many variables that affect costs in the forecasting process, forecaster should be able to evaluate what kind of forecast can be executed with what kind of costs. Budget should be kept in mind during BI&A process especially when creating forecast.

Collecting and processing the date is the initial step of categorizing the data according to the target of BI&A process. As stated in the chapter 3.1.3 in the forecasting process, cleansing the data from background noise is the initial step to ensure representativeness of the data. This requires understanding the data and collecting additional inputs from business to identify possible wrong or biased data entries.

3.5 Requirements for the data in analytic process

Data source should be considered carefully in business intelligence and analytics (BI&A). As stated earlier in the chapter 3.1.4 forecasting can be divided into quantitative and qualitative forecasting, based on the used data source. Pirttimäki & Hannula (2004) note that only one information source is not generally enough for effective decision making and both data sources should be considered. In this case information is used as a term for all levels of understandings. Third category that should be considered is the subject of the information. They have presented these categories as intelligence cube which describes information source, subject and type which. This categorization is shown in the figure 7.

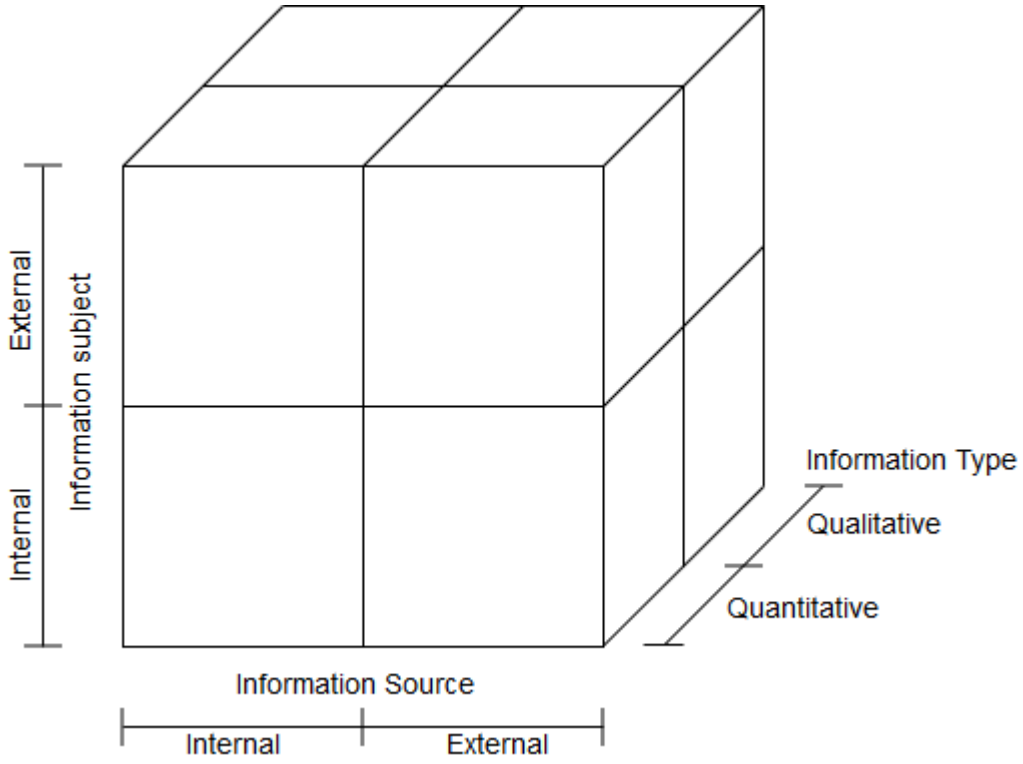

Figure 7. *Information cube (Hannula & Pirttimäki 2004)*

Information cube provides reminder for BI&A process which aspects of data should be considered. For example, organization's employees have often lot of information regarding competitors. This would be qualitative, internal information about external subjects. (Tyrväinen 2013) In the scope of forecasting, data source, type and subject are all needed to be considered in the beginning of BI&A process. In the first step subject of the forecast needs to be considered: does the forecast aim to produce intelligence of internal or external factors. Considering the target of the forecasts should be in line with the business need.

Armstrong (2001) has evaluated different data sources that are commonly used with statistical forecasting methods. Data needs to represent the events being forecasted. Historical data is often used as the basic data source for forecasting (Stevenson 2014). As stated in the chapter 3.1.2 there are situations where historical data can't be used. This might be due market situation or in the case of new products for example, there might not even be historical data. Data for statistical methods can also be obtained by finding data of comparable situation or by conducting experiments (Armstrong 2001). Table 4 presents comparability of different data sources to different forecasting purposes. Suitability is evaluated from 1 to 4, where 1 is most suitable and 4 is least suitable.

Data Source	To reduce costs of forecasting	To control effect of researcher's bias	To estimate current status	To forecast effects of small changes	To forecast effects of large changes
Historical	1	1	1	1	4
Comparable situation	2	2	2	4	4
Laboratory experiment	3	4	4	3	2
Field experiment	4	3	3	2	1

Table 4. *Suitability of data sources to forecasting purposes (Armstrong 2001)*

As seen from Table 4, historical data is suitable source for many forecasting situations. In the case of forecasting large changes, it should not be used as only forecasting method. That would require more information about the future and experiment would be most appropriate method to obtain data. On the other hand, historical data could still be used to validate the forecast done from data obtained in some other way. As stated in chapter 3.1.4 it is advantageous to use multiple forecasting methods to validate the forecast.

Importance of data is can't be underestimated during the whole BI&A process, but it is emphasized especially in the first four steps of the process. Definition and business need determine what data sources should be used. In the light of forecasting this means that they also determine the appropriate method for the forecasting, as the data source is the primary variable which determines whether quantitative or qualitative forecasting methos should be used (see chapter 3.1.4). When used data source is defined, this guides the forecaster towards collecting sufficient amount of data and using the correct methods for this data.

As seen in the Table 4, in the case of using quantitative forecasting methods with historical data as the source, the amount of data is also variable which affects the forecasting process and methods. As stated in the chapter 3.3. business need acts as the trigger for BI&A process and it should be kept in mind during the whole process. When conducting forecasting using the BI&A process as the framework this means data collecting should be done based on the data definitions. Therefore, during the data collection step, it should be considered whether the business need of the forecast can be met with available data, or whether some adjustments to either the budget or the accuracy of the forecast should be made. This requires reevaluation of the business needs and expectations

4. EMPIRICAL RESULTS

4.1 Interviewees' backgrounds

As noted in the theoretical framework understanding the underlying business needs is the starting point for BI&A process as well as an important factor of forecasting. On the other hand, it was as well noted that organization might have trouble with identifying the information needs. For this reason, background of the interviewees was surveyed in the interviews. Most of the interviewees had worked for target organization for several years and were therefore very familiar with the organization's culture and processes. Importance of tacit knowledge gained from experience from working for the organization or its field in general was theme that rose up during interviews.

I2: *"We can identify the standard products, because of gained experience. System doesn't know which products are standard and which are not."*

I7: *Understanding of the market situation and pricing comes from sales department. We produce baseline price for product which defines the low-limit for pricing, but sales department define the price for the product based on the market situation.*

I10: *"There are a lot of exceptions in sales data. These exceptions don't show anyway in NetSuite and the analyst just needs to know them."*

Roles of the interviewees in the target organization are presented in the Table 1. Target organization is most importantly a growth company as noted in the introduction. This has led to situation where employees get to participate in various areas of business and this way gain a lot of tacit knowledge of the organization.

I2: *"I work as distributor business manager. At the same time, I'm taking care of big level business analytics for sales operations for all the group levels and it is only based on the market-based pricing."*

I7: *"Currently I'm working 70% of my time as product line manager for our Top Hammer equipment. In addition to this, I work rest of my time with marketing and our digital Sense-product."*

4.2 Analytics currently utilized

As noted in the theory, there are various form of analytics. As a term, analytics was understood differently depending on the persons position in the organization and their usage of analytics. Generally, definition for analytics was asked to define based on the analytics interviewees themselves use in their day to day work.

I2: "*Organization culture is not analytics based. Maturity of know-how with data analytics and information systems in general is on very low stage with the sales personnel globally. Top management have not also used data analytics themselves and don't necessarily see the value in it."*

I4: "*Generally, I run various reports from ERP system. I use them to analyze and monitor order and subcontracting situations. Also, delivery schedules are monitored."*

I5: *"I'm responsible for sales reporting and monitoring as well as budget monitoring done for upper management. I create material for Sales and Operations meetings and have created the process used in them."*

I10: *"My task has been inventory planning. I analyze inventory status, not just in Finland but globally in all locations to suggest stock levels or determine what tare the slow-moving items and how do deal with them."*

From technological point of view, all interviewees recognized Microsoft Excel as the main tool used to execute analytics and reporting. In the analytically themed interviews, it was also discussed whether interviewees are familiar with business intelligence tools such as Microsoft Power Bi. Since the focus of this thesis and the interviews were not in the technical solutions, Power BI was used only as example because target organization has Enterprise level subscription to Microsoft Office 365 which would allow using basic functionalities of Microsoft Power BI without additional costs. Much interest regarding potential of BI tools was expressed by the interviewees but it was also noted, that effective use would require training.

I6: *"In the next 0.5-1 year we should not invest in any new system. These don't take us forward but rather take our attention away from the basics. We need to build strong enough foundation, which is scalable and only after this, implementation of new systems should be considered."*

Descriptive analytics seemed to rise as mostly used form of analytics. More analytically oriented personnel were also, able to additional types of analytics currently utilized. Various functions in the organization were identified to execute also predictive analytics in the form of forecasting. Understanding of forecasting was also divided depending on the persons position in the organization. Financial function saw the forecasting especially as strategic tool whereas supply chain function emphasized the importance of operative demand forecasting.

I3: *"Demand forecasting is information of future sales and customer requirements. It is also understanding how we can affect demand with pricing solutions and channeling demand. Sales function does demand forecasting. They use monthly sales & operations meetings to identify future sales, customer leads and to evaluate statuses of current projects.*

I5: *"We create forecast of profit and loss on the concern level. This is done on quite rough level by taking realized sales to date and collecting data of future sales from market area vice presidents to evaluate how different market areas are going to develop. This forecast is communicated for board of directors, chief executive officer and chief financial officers. If needed, forecast of profit and loss is also communicated externally according to stock exchange principals."*

I6: *"We have long delivery times because heavy products are moved as sea freight. We would need to be able to better anticipate where which products have demand and store appropriately."*

Analytics and forecasting in the target organization generally been based on historical data. Analytics and forecasting have been executed in ad-hoc basis based on business need. No clear structure for analytic process has been created. This has been seen as reality of the business especially since the demand is often much project based. At least half of the demand was described to for projects.

I4: *"More you know, more you prepare and understand. It needs to be accepted that we are in the business of consumable products and the biggest orders are for project-based. Interfaces need to be accepted and forecasting is difficult. Long contract projects ease forecasting and this operating model system seconds."*

4.3 Forecasting targets and opportunities

Purposes of using forecasting are divided between producing intelligence for strategic and operative decision making. Strategic level decisions should be in line with organization's long-term strategic targets to which forecasting should provide insight.

I2: *"Stable products need to be defined to create efficient forecasts."*

I5: *"Budgeting sales for is done yearly. Targets in formed budget are in line with Robit's long-term strategic goals. For this year vision was 15% organic growth based on the strategy. This created basis for the country-based budgeting. Currently budget is monitored on product line level in a country."*

I10: *"Forecasting is based mainly on the historical data and meetings with sales manager to define stock level. It is used for inventory planning. Every six months this forecast needs to be set again depending on the new and ending projects."*

The scope of this thesis is to focus on forecasting on more operative level. As stated in the theoretical framework strategic level forecasting is often biased because it is done based on the organization's target rather than historical data and other unbiased data. An-

alyzing the inventory levels reappeared as a theme with all analytically themed interviews. Sales of target organization's products is divided all trough world and when this is combined with long delivery times this creates problems for warehousing.

I6: *"Sales could be increased with faster deliveries which could be achieved with warehousing right product at right place. Much consideration regarding products should be exercised when choosing stockable products so the inventory value does not rise."*

I9: *"Robit's will is to sell through dealers, who should hold products at their warehouses. For the dealer to warehouse, they would require that also Robit has additional warehouse close by from which they could resupply."*

I10: *"I also analyze whether slow moving items could be sold somewhere else. I create list of slow- and non-moving items which sales department double check: whether it is not moving for a reason. If item has potential for sales in normal price we remove it from the list. The final list is used for campaigns or something to get items moving."*

Products which are going to be warehoused should, by this logic, be forecasted and analyzed. As theoretical framework suggests, forecasting on higher level produces more accurate results easier. Therefore, forecasting could be developed on country or market area scale since the target organization exports their products all over the world.

4.4 Demand types

As noted in the introduction scope of this thesis is to focus especially on continuous demand. From the beginning of the research it was clear, that demand of target organization is split between continuous demand and project demand. Forecasting stable products with continuous demand is good starting point for forecasting, since in this case, future often is reflection of the historical data. This allows taping easily into sales data from ERP system to create forecasts. Problematically this would also require that demand types are somehow separable from each other.

I1: *"Continuous and project sales orders are not recorded differently in NetSuite"*

I2: *"Reality of our business is that historical data is not good basis for forecasting since most of our business is based on project basis. In projects specific item is needed which we manufacture and when the project is over the same item might not be sold in coming three or four years."*

I4: *"As far as I understand it, NetSuite is not capable to separate whether product has gone for project- or continuous demand. This information is in the head of persons associated with the project."*

I7: *"Product categories can allocate whether product is meant for continuous or project demand. Piling-products are basically all for projects and in case of Top Hammer products are often continuous demand. Especially smaller sized drill bits are often used in mines, which are continuous demand."*

I9: *"Mines which are in continuous demand can be identified. Small-sized drill bits are continuous demand. For example, 43mm, 43mm, 48mm and 51mm smaller bits are generally continuous demand. Predictability of project demand is very difficult since the timeframe in which product is needed is often very tight and the one who has appropriate product in stock is usually the one who gets project."*

Identifying the demand type was noted to be one of the most difficult tasks when considering forecasting. Employees who had been working for target organization seemed to have the general understanding of which product is for continuous demand and which is to project demand. Unfortunately identifying these from analytical point of view proved to be very challenging when asked in the interviews from employees who execute analytics.

I3: *"Historical data can't be hold accurate. New products are launched often and the information about future sales is on the field. Sales knows what their customers are going to buy better than what NetSuite's historical data would suggest."*

4.5 Communicating analytics

High inventory values rose up as a topic in many interviews. Historically target organization has had problems with identifying demand types and especially communicating signals of changing demand. This has led to increased inventory values on global scale with slow- and nonmoving products causing biggest problems. In analytically themed interviews analytics of these items was also discussed and the target organization has lately been especially trying to identify whether these slow- and nonmoving items would have demand on some other location. Missing visibility of the warehouses and incorrect inventory levels have caused problems with these analytics.

I6: *"In the light of history, we have created products needed for projects to stock and when the project has items have remained in the stock. Our system has capabilities of inventory planning but currently we are not ready to start using them. Data quality would need to be increased."*

Especially when long project is ending, signal should be communicated to all relevant parties. As noted this has not been the case, but target organization has tried to resolve this situation with sales and operations planning (SOP) -meetings. World has been di-

vided into eight sales regions in the target organization and each region holds SOP-meetings monthly. In interview I3 it was noted that SOP meetings utilize both historical and judgmental data. Sale personnel produce judgmental data of coming sales and the overview of the market and additionally financial functions provide SOP meetings with historical data of the sales.

I2: *"Sales and operations planning workshops are important for the forecasting of coming projects. Start of the project is identified so manufacturing knows when they need to have products ready."*

I5: *"In practice our demand is very hardly forecastable because majority of our customers are project demand and smaller continuously buying customers. Based on Robit's strategy we also execute much of our sale through dealers whose sales we have no clear visibility. In the SOP process, much time is spent especially in the evaluating coming sales."*

In sales themed interviews target organization's selling strategy rose up as a theme. Strategy has been to increase sales through distributors or dealers rather than selling straight to the customer. This has been the strategy of target organization, since currently company does not simply have the resources to have a sales office in every corner of the world. Dealers also often have good relationships with their customers, end users of the products, and this allows target organization to get into various markets to which they would have no access without the dealer.

I5: *"We aim to sell as much as possible through our dealer network. Our own sales organization is very slim, and more time sales employees spend taking data to NetSuite, less time they have to sell."*

I7: *"For forecasting, we would need input from our dealers. Currently about 40-50% of our sales comes through dealers. This portion is targeted to grow in the future to 60-70%."*

I9: *"We don't produce forecast of coming year to our dealers. Dealers don't want to tell us how they see sales of product group in the coming year. We would need to get systematically more information from the dealers. Currently we need to milk much information from them and document it in various places."*

Communicating data was noted as problem with the interviews. Both internal and external communications seemed to have problems. Internally forecasts of coming sales did not seem to be communicated efficiently to relevant parties such as purchasing and subcontracting. Communication with external stakeholders was also noted to be highly depended on the individual responsible for it.

I1: *"No forecast is given to our subcontractors and we don't get any forecast"*

I4: *"We don't deliver forecast or our view of the market to our subcontractors. In some cases, demand delivery times are evaluated together with the subcontractor."*

I7: *"With our support, sales need to create structures which would allow which would allow forecasting demand based on the demand that comes to dealer. Communication between sales should come to offering and sales support -function better than currently"*

Systematically done external and internal information sharing was noted to be missing in many interviews. Communication is currently highly depended on individual and whether they share their information. Lack of formal communication rose up in all functions which were interviewed. Even though SOP process offers a link in between the sales and production, the offering personnel are not included in these meetings.

"Process for internal analytics is needed. More frequent reviews and more systematically build of stock levels. After internal structure has been created, additional external data sources can also be looked at to get the idea what kind of growth we can realistically expect."

In the case of forecasting this systematic approach would mean identifying the important variables affecting demand. Based on the input from sales, offering and sales support -function manages the product portfolio to match demand with appropriate products. This would require also additional input from other functions.

I7: *"Supply chain and production input requires development. Communication is done on personal level, but it is not adequately formal."*

4.6 Data

Especially in the analytically themed interviews the role of the data was discussed thoroughly. As noted in the theoretical framework, forecasting requires that there is some data available that reflects the future. During the interviews it became clear, that in the target organization's case, this means identifying various data sources about the demand, understanding much about the business and products. Understanding the basics of analytics is of course needed, judgment as the data source was emphasized as more reliable than using solely historical data as a source.

I2: *"Previous sales indicate future sales but at the same time we need to talk with the customer as well."*

I4: *"Data in NetSuite mirrors our current situation. System does not give a view of the future. Forecasting tomorrow, based on the historical data is challenging. You need to know isolated nuances for example from single project products, which's demand is going to end so that you would have condition to succeed."*

I10: *"Historical data does not have information whether it reflects future demand. This information is tied to the sales manager."*

Understanding how data is recorded in the ERP system was also repeating theme of analytically themed interviews. In the interview I3 it was noted that some forecasts are done for customer. From the system's point of view this proved to be difficult since customers are created under subsidiaries. This means that single customer has multiple customer ID's in the system if they have purchased products from multiple subsidiaries and analyst needs to be aware of this feature.

I2: *"Customers are created with different ID's for different subsidiaries. System can't combine these different ID's to get the total sales of the customer. Currently, this must be done manually. If there is a difference in customer's name, it is even harder since Excel don' recognize these."*

4.6.1 Sources

Data sources were identified as important part of business intelligence & analytics framework in the theoretical framework. Especially when creating forecast data behind it must be representative of the phenomena being forecasted. Therefore, each interview included a question about used data sources. This allowed to create broad understanding of data sources used in reporting, analytics and forecasting in various functions of the organization. Target organization's ERP system was noted as the main source of data in every interview. It contains historical data about sales, customers, subcontractors and so on. Especially the historical sales data of invoiced orders was seen useful in the scope of forecasting.

I1: *"Invoicing is the best metric for realized demand."*

As Robit uses Office 365 globally which allows sharing files and information easily in the whole organizations either with usage of OneDrive or SharePoint it natural that employees take usage out of the available systems (Skendzic & Kovacic 2012). Challenge regarding with this is though, that since the ERP system was seen as the main source of data in every interview, it should be the place with most up to date data. Unfortunately, empirical findings from interviews reference that this is not the case.

I10: *"In Robit, employees are still working with many different Excel files and they haven't updated the latest information to NetSuite. Hopefully, in the future we would use NetSuite more efficiently and get up to date information straight from there."*

The existence of additional spreadsheets and other data sources was recognized in all interviews and in all functions that were interviewed. In each interview various data sources were identified. Problematically different functions did not know about data

sources used by other functions. For example, financial function even had a business intelligence system which could be used in forecasting but supply chain function did not have access or even knowledge of the system. Many interviewees answered also that in the target organization many processes could be done more efficiently with the help of existing systems and if the systems that is used for each process would be clearly defined so that for example, data about future sales would not need to be filled to multiple systems.

I2: *"There are future leads in the Excel file shared with SOP planning team. It contains information about new projects and possible demand. There is also information of the probability on whether we get the deal or not."*

I4: *"KPI's of the factory are reported in Excel file in SharePoint. Our steal consumption is partly in NetSuite and partly in M-Files depending on how long the steal has been on our stock."*

I5: *"Our bookkeeping is not currently in NetSuite. IBM Cognos Controller is used to consolidate our subsidiaries bookkeeping from various external sources. This data is also linked to IBM Planning Analytics which is new additional system used in strategic planning, budgeting and forecasting. Integration with NetSuite data is not integrated to these systems, but it is a possibility in the future."*

I6: *"I have understood that the sales department have access to external data sources, which compile information about projects where our products could be used. It is used as a way to gain market intelligence. Especially mining and infrastructure projects can be perceived from them"*

I9: *"We have a service which send sales department weekly emails about upcoming projects on global scale. Even though it most likely costs Robit much, it is not used effectively. In addition, each sales manager has their own knowledge sources about projects going on in their area. In SOP meetings, biggest upcoming projects are looked into."*

All functions noted to use other data sources which contains information that affect demand and forecasting. In addition, all interviewees were asked what external drivers affect the demand of the target organization. Many external drivers were identified during the interviews, but the effect of these to target organization's demand can't be reliably estimated in the scope of this thesis since the business area is affected by many variables which can't be controlled. Especially timeframe is was hard to estimate when asked how and when these variables affect target organization's demand. Many of these external drivers such as drilling rigs, were seen to affect the demand, but it was also noted especially in the sales themed interviews in these cases affect to target organization's demand would realize after maybe two years and current forecasting capabilities don't enable reliable forecasts in that long-time horizon.

I6: *"Construction business is a big driver. Also, the prices of raw materials affect from how difficult places it is profitable for our customers to drill for them. Of course, the prices are same for all."*

I7: *"Development of investment products in our field affects our demand. For example, couple years back, our Evolution product family was developed because impact power and frequency of drilling hammers in underground mining industry."*

I9: *"Possible coming EU funding, in the field would affect our demand."*

4.6.2 Quality

Quality of data was vastly covered topic during the interviews since it was identified in often as the main source of data used in analytics and forecasting. Quality of data referenced to the data stored in target organization's ERP system as it was identified as only consistent data source throughout the organization. Depending on the interviewees, role in the organization many different views of the data was considered. From single function's point of view data quality was not seen as a problem whereas from more overall point of view

I2: *"Historical data is correct in NetSuite since it is based on invoice which always goes through the system"*

I3: *"Warehouses are in NetSuite where they have stock levels. The problem is that warehouses have many storage places and products. Inventories have not been reliable. Subsidiaries might have had additional more up to date spreadsheets which have not been updated to NetSuite and therefore data in NetSuite does not add up. Also, visibility to warehouses have been problematic"*

I10: *"There are problems with historical data, since some tacit knowledge is not updated to NetSuite and future sales might actually differ much from the historical data."*

As noted many challenges regarding data quality have been identified and therefore, historical data can't be completely trusted when creating forecasts or basing other analytics on it. Problems with data quality have been noticed by ERP development and steps towards correcting data quality have been made. Challenges also lie in understanding the data and how it is generated.

I2: *"The problem with data is that if customer buys from our subsidiary internal transfers should be removed from the data. For example, we can have a customer which buys product from our subsidiary in America. America then makes a sales order from group or production site so the sales order from customer can be completed. This creates a duplicate and internal sales should be removed."*

I6: *"We have cleansed order data. Meaning going through order masses, closed old orders, validated recent order and sales data, validated inventories and added internal and external sales prices to the ERP system."*

As noted in the interviews, understanding and fixing the problems with current data is half of the solution. In addition to this, the other half is creating effective processes that ensure collection of high-quality data. Otherwise, eventually the data quality will again decline, and more cleaning efforts are needed. Interviews I3 and I6 noted that target organization is still on the process of bringing all their subsidiaries in the same ERP systems as the rest. In both interviews, also the importance of creating right processes to ensure that all data goes through ERP system was emphasized. The importance of this was highlighted since currently many subsidiaries don't have historical data available. When systems are used accordingly and data collection correctly, this ensures that in the future data quality will iteratively improve when new data is higher quality and older low-quality data becomes more irrelevant. Challengingly this means that organization culture needs to change to ensure that new processes are followed.

I6: *"From my part, I also have to monitor that Korea's inventory would start to move. We have identified reasons why product movement has been slow and agreed about correcting actions. Now monitoring is needed to ensure that situation straightens, and we don't return to old habits."*

Understanding the data and its connection to the business need behind BI&A was acknowledged in the theory. Empirical findings seem to also endorse this since understanding the business of the target organization rose as important theme in many interviews and in the analytically themed interviews the importance of running reports with correct attributes was emphasized.

4.7 Product portfolio

Product portfolio regarding Top Hammer drill bits rose up as repeating theme especially in the interviews with manufacturing, offering and sales. Target organization has vast product portfolio for Top Hammer Drilling products from which especially drill bits were the focus of this research. Robit's Lempäälä manufacturing plant focuses especially on these drill bits and there are most of historical data available from them. From forecasting point of view this creates problems when customers can order from vast product selection various items.

I4: *"Customers can order even single product for the same price as one. Grouping multiple orders of same product is unfortunately not also option for manufacturing because the probability of order for the same product in near future is often low. We have very vast product portfolio and should consider screening product labels if possible."*

From the sales point of view, large product portfolio was seen as sales advantage as long as sales personnel understand the usage for each different product. For production this creates problems when often small quantities of specific product are sold. From supply chain point of view this requires definition to analytics to define what products have stable demand and what products should be forecasted.

I1: *"Optimizing batch sizes in production would be important to follow. This would allow to create minimum order values."*

I6: *"We have aimed to categorize products, but this has not been done systematically. Products categorization process needs improvement. Currently we speak of runners, strangers and repeaters."*

I10: *"Currently we have a rule that if during a 12-month period item has been purchased 3 or more times, this item is considered as runner which means it is stockable. From these item lists is created which is validated with sales to double check. Whether project is starting or ending this might affect."*

Extension of product portfolio was noted as one of the most challenging variables in the forecasting context. As the categorization can't be trusted as the implication for forecastable products other themes rose up during the interview. Grouping products based on their attributes rose up as theme which would help forecasting Top Hammer button bits. Variety of the products rose up many times during the interviews, but in no specific product was mentioned to be unnecessary.

I4: *"We are in project-based consumable product business which makes forecasting very challenging. I would see that this problem could be controlled with warehouses. There can't be too much warehouses though. In our situation, variety of product portfolio complicates both warehousing and forecasting. Only in Lempäälä we have several hundred products which in rotation."*

I7: *"Basic determinants are thread, diameter of the product, different face models of the button bits, used blank and form of the bits. Especially diameter of the product is important for the customer. Often, we can't suggest switching of the drill bit's diameter. Different diameters are needed to be available. Inside single diameter though, we have too many face models for example."*

Guidance of project portfolio and by extension, demand was asked in the interviews. In many interviews it was noted that slow moving items have been problematic and raised values of inventories. Only systematically done guidance of products was noted to be yearly organized sales campaigns of slow- and non-moving items. Price-based demand guidance was noted to be possible area of improvement. In the interview I7 it was noted that input from supply chain and manufacturing would be needed to develop minimum order sizes which could be used to guide demand.

I7: *"We offer trainings to our own personnel and to our dealers about our products. With these we aim to present our product portfolio and usages of each product."*

New products affect the project portfolio. In the interview I7, new product development was discussed. Generally, in the case of Top Hammer products, two new product inquiries come every week. Based on the supposed order size and strategic reasons determine whether new product is developed or not. These requests are generally minor modification to existing products and their development doesn't take that long but on the other hand this creates additional product for the product portfolio.

I4: *"New products to replace existing products are don't come often. New product models come quite often but when these are published, old models are not killed, and they live on."*

I7: *"We have been often flexible with acceptance of new product development requests because we have wanted to be agile. Often we have made new product rather than tried to offer almost similar existing one."*

4.8 Delivering to customers

As public limited liability company, the target organizations main goal is to produce profit for the owners of the company. As noted in most interviews, target organization works in the consumable product project business. In this situation, ability to deliver products is essential. Even though different functions inside the organization emphasized a bit different targets for the forecasting, the target of forecasting is still the same: to be able to respond to demand in appropriate timeframe.

As noted earlier demand is either continuous or it is for projects. Repeated theme in the interviews was that customers generally want their products as soon as possible. This requires that organization's supply capabilities are adequate.

I2: *"We don't have a culture to create ideas because we have so many things to do and are scared the implementation fails and the time is wasted. It would be better to enable idea generation and implementation since we would either way learn from these projects."*

4.8.1 Sales requirements from forecasting

Originally sales input regarding forecasting was intended to be let out of the scope of the interviews. As the interviews progressed, and after a suggestion by an interviewee, additional sales representative was also interviewed. From all interviews the main finding was that sales process is very vague and information about upcoming sales is hard to tap into as an analyst.

I2: *"Nobody has used NetSuite's leads capability. Offer is demand which you either get or don't get. All offers supposed to be made through NetSuite"*

I9: *"There have not been many trainings regarding NetSuite. Making an offer, creating sales orders and checking inventory levels have been taught. If CRM capabilities would be introduced, additional training would be needed to ensure that process creates data accordingly."*

During the interview it came clear that target organization's ERP system contains also many CRM capabilities should as leads, opportunities and customer activity recordings. Unfortunately, from analytics point of view, these are not utilized effectively and therefore, do not offer additional data for forecasting. Interviewees I2 and I9 are both working in sales. They were asked to describe typical sales process which is presented in the Figure 8. The process has three potential starting points which are marked with triangle.

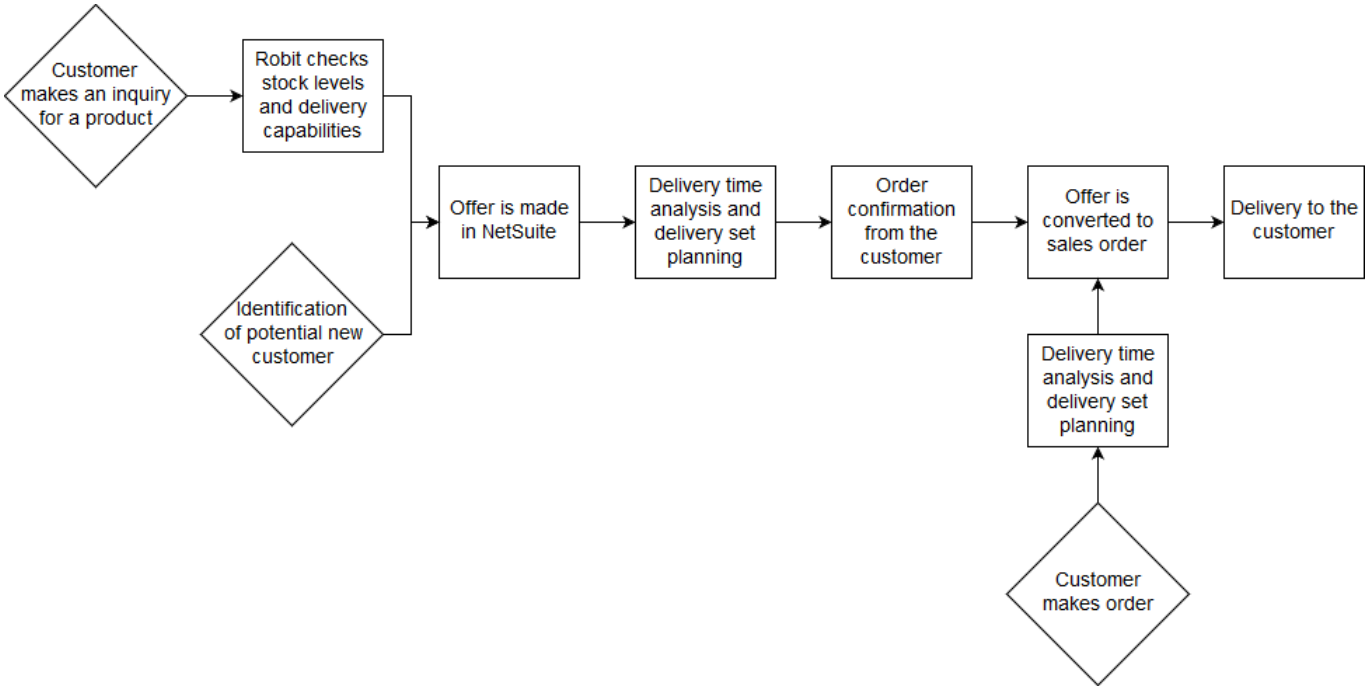

Figure 8. *Typical sales process at the target organization*

Through the sales process, currently data is registered in ERP system only when the offer is made, when offer is converted to the sales order and naturally regarding deliveries and invoicing. As noted earlier, there is data regarding the sales opportunities and leads in the monthly sales and operations planning meetings. Basically, this means that the first data about possible client is only in ERP-system when the offer is created. Customer's use of the ordered product is not recorded anywhere. This means that the information whether order is for continuous or project demand is only known for the sales representative responsible for the customer.

I9: *"Modification for the offer is not made in NetSuite. The negotiation is done in other channels. Also, we don't evaluate currently the probability of getting the offer, when it is made even though NetSuite would offer possibilities for this."*

Sales offers were noted as a possible source of information. As also noted, probability of target organization getting the offer is mainly discussed in the SOP-meetings. Understanding how many deals target organization is going to get from the offers would offer an additional source for data when forecasting. In shorter time horizon open sales orders offer view to near coming future.

I5: *"Our offering quantity is relatively low and from it is hard to forecast demand. Current sales order quantity equals roughly sales of one month."*

4.8.2 Supply chain requirements for forecasting

From the supply chain point of view the importance of having the right product at right place at right time was repeating theme in the interviews. As noted earlier target organization has long delivery times when heavy products are moved as sea freight. During the interviews it was noted that faster deliveries are sometimes done with airfreight, but this is used mainly when swift deliveries are needed in smaller quantities because of the higher cost of delivery.

I6: *"If you don't warehouse product, it should not be forecasted. This does not hold completely since we need to have raw materials in stock for products which are produced based on order."*

I10: *"Long supply chain from production to customer. Sooner we know the demand the better and therefore more efficient forecasting should be used. We should also deliver the order in partitions if possible."*

Warehouses around the world were noted as the main enabler of swifter deliveries. Target organization holds their own warehouses globally and additionally dealers were noted to hold warehouses. As noted earlier in the empirical framework, visibility to warehouses was noted as a challenge. In interviews I3 and I10 it was noted that warehouse stock levels are being analyzed and compared to global demand. From the supply chain point of view, also the current lead times are considered in the manufacturing plants. Based on these factors it is analyzed how long is the lead time from order to delivery.

I3: *"Forecasts are used for stock levels. What we have too much in stock and whether this could be sold somewhere else. Forecasting has focused on optimizing the stock levels, in which we have huge problems. These problems have not been yet corrected."*

I10: *"We should focus more on internal analytics currently. We have any many internal analytical problems for example with supply chain which should be solved first. This would create good basis to start doing external analytics as well."*

Regarding continuous demand, supply chain's most fundamental need for forecasting is to know what products to warehouse and where. Because of the variety of products and long delivery times this is challenging. Additionally, data regarding future demand was found out to be very scattered throughout different functions.

I5: *"If forecast would be on product level, this could be used in production planning, but we are very far from this. Data in NetSuite is no by itself enough to do this."*

4.8.3 Manufacturing requirements for forecasting

Attributes affecting manufacturing were considered mainly from the demand point of view. It was noted in the interviews that target organization's production is triggered by either upcoming projects where products are made based on the order or by alarm limits of stocked products where products are made when the quantity of specific product in warehouse reaches its minimal value.

I1: *"In production we would like to know what steal is required. Steal has currently 3-4 months delivery time so forecasting time horizon should be 6 months and regarding products we would need to know the diameters to be required. Forecast of single products doesn't create that much value for production."*

I4: *"For production, forecasting demand means that we have right conditions in place. In this case with Top Hammer products this means that we have appropriate steal and bits available so that we can carry out any project within our production capability."*

5. DISCUSSION AND CONCLUSION

The first sub target of this research was to provide target organization with understanding on what data sources are required for efficient demand forecasting process. During the conducted interviews, it was noted that different functions in the target organization have different requirements for the forecasting. These requirements have been compiled in the Table 5, with data sources associated with forecasting requirements.

Function	Forecasting need & scope	Required data sources for the forecast currently
Supply Chain	Forecasting demand and optimizing warehouse stock levels based on this. Operative forecasting.	- Historical data from NetSuite on item level including use of the product - Judgmental input from SOP-meetings on item level of upcoming sales which are not in NetSuite
Sales	Forecasting for upcoming sales. Information about delivery capabilities. Collecting information about future leads and opportunities. Operative forecasting	- Information about coming sales from sales managers and their distributors. - Possibility of using leads and opportunities. - Delivery capabilities based on stock levels and current production situation.
Offering & Sales support	Forecasting product portfolio. Understanding coming requirements for products and manufacturing capabilities of these. Guiding offering based on this information. Strategic and operative forecasting	- Historical data from NetSuite on item level - Information from production about optimal production batch sizes - Information from sales about product needs of the customers
Manufacturing	Forecasting required for raw materials. Especially forecast of steal consumption would be helpful since delivery times of it are long. Operative forecasting	- Historical sales data from NetSuite on item level including diameter of the product. - Consumed steal data from NetSuite and additional Excel file. - Judgmental input from sub-contractors about consumption.
Financial	Forecasting done for budgeting purposes according to Robit's strategy generally on group-level. Strategic forecasting.	- Historical data from ERP system consolidated on country and product line level - Judgmental input from Sales function consolidated on market area level.

Table 5. *Forecasting requirements of different functions in the target organization*

Forecasting needs in the target organization were divided into operative and strategic forecasting. Strategic level forecasting was especially done in the financial function in the forms of budgeting and overall sales development. This form of forecasting is done based on historical data and judgmental data from various sources and used in reports for upper management and board of directors. Data used in these analytics is compiled overall and therefore, offers view on high level without many details. This also leads to higher accuracy of the forecasts than more detailed forecast would (see chapter 3.1.1).

5.1 Requirements for the data

Operative forecasting focuses on more detailed level. Most of the functions identified needs regarding operative forecasting but also noted that in the current situation, it is very hard in the case of the target organization. During the interviews it was noted that reality of the business is that most of target organization's sales comes from either projects, or through dealers. This affects forecasting, since it requires data, that provides understanding or insights of the future (see chapter 3.1.1). In the case of continuous demand, this is often the case and the customer have stable demand, but projects were noted to happen often on short time period and important thing was that when they begin, delivery capabilities of the production and warehouses are adequate. The data regarding upcoming sales was noted to be very disperse. Sales which comes through dealers were also noted as challenging to forecast, since target organization does not provide forecasts generally for their dealers and they don't require forecasts from the dealers.

Communication between target organization and their dealers is highly depended on the target organization's sales department and it was noted that more information would be required from the dealers. As suggested in the chapter 3.1.3, sharing forecasts with relevant stakeholders, in this case the dealers, would provide mutual value. Dealers would know what target organization expects from them and how much they products they have analyzed can be sold through dealers. Target organization could require the dealers to correct the forecast if the dealer's view of the future sales differs which would provide additional input about future sales and increase cooperation with the dealers.

In addition to increasing external communication with the dealers and other stakeholders such as subcontractors in production, to obtain new data, internal information sharing should be increased. In the chapter 4.6.1 it was noted that target organization has various Excel-files which contain more up to date information than ERP system. For example, Sales and Operations planning (SOP) meetings were noted to have much information from upcoming sales which is currently in no other location.

In the defined analytics framework, it was noted that for efficient analytics, access to relevant data is required (see chapter 3.3.). Currently operative forecasting of demand has been based on NetSuite's historical data. Forecast has been mainly used as conversation opening rather than basis of decision. Including judgmental input from SOP-meetings and adjusting the historical forecast based on that would increase the accuracy of the forecast when two data sources and forecasting methods are used (see chapter 3.1.4).

As noted in the interviews, internal communication is currently mainly executed on individual level. SOP meetings are an exception in this since they aim to offer transparency of the upcoming sales to other functions. These meetings are hold monthly for each sales region and contain much information about future sales which are in separate Excel files shared with relevant parties. Moving this process to more accessible platform such as the

ERP system would allow analytics of this data much more efficiently. Noted in the interview I5, the SOP process was currently considered to be moved to the new IBM Planning Analytics -solution, though no other function but the finance seemed to be aware of this.

There were multiple findings about the data in NetSuite. As the sales process and invoicing go through the system, invoicing seemed to be the best meter to evaluate historical demand. Future demand could also be evaluated trough sales orders, but they only provide about 1-month view of the future sales (see chapter 4.8.1). Importantly from the forecasting point steps to ensure data quality would be required.

Manufacturing noted that forecasting future demand by grouping sold products by their diameter size would help to forecast steal consumption which is used as raw material. Consumption of raw materials need to be forecasted to be able to answer to demand even in the case of products which warehoused (see chapter 4.8.2). To evaluate possibility of this, report from NetSuite was run about the sold Top Hammer button bits on a product level with diameter included. Only roughly 15% of items had the information about the diameter. By enrichening the data this could be increased to roughly 40% but still much of the items were missing the information. Robit's public website contains the information about all Top Hammer product diameters (Robit Plc). Since this metadata for the product is already available this information should also be added to the NetSuite to improve amount of data available to forecast diameters.

It was brought up many times during the interviews, that project demand and continuous demand are not separable in the ERP system. Even though, with deep understanding of the products, it was noted to be possible to separate project demand from the continuous demand, empirical part did not offer clear insight and much more understanding of the products would be needed. Even two demand types could be separated based on product, this wouldn't be a solution in case of long projects to which products are made into warehouses. Important information would be to know when the project is going to end so no extra products would be made and warehoused for ending project. (see chapter 4.4) To resolve this problem, additional information about the sales orders would be needed about the customers use of the product. Simply, this could be done by including demand type as a mandatory field to new sales orders. Demand would either be continuous which is seen to last indefinitely or project demand which is estimated to last until specified date. It would help if this information was in NetSuite and would need to be separately collected (see chapter 4.8.2).

5.2 Requirements for the data processors

Second sub target of this research was to provide target organization with understanding what is required from the data processors when forecasting demand. As noted in the interviews, current data is not by itself enough to forecast solely based on historical data. In addition, many of target organization's subsidiaries were just introduced to the ERP-

system and therefore don't have much historical data available. Therefore, on global scale forecasting based on historical data can't be hold reliable and much weight much placed on judgmental forecasting. Theoretical framework seconds this notion, since advanced forecasting methods which are based on historical data, often require at least three years of data (see chapter 3.1.4).

Maturity of the organization's analytical capabilities was identified to be low (see chapter 4.1) and from the forecasting point of view, their demand is challenging to estimate because of the projects of which there are little data available. To increase reliability of the historical data forecaster needs to have access to judgmental data sources. Sales were noted to have the best perception of the future demand which is evaluated in their SOP meetings. Therefore, forecaster should interact closely with sales department and participate all SOP meeting to get the latest information about future sales as well as ending projects. As the SOP meetings were not noted as the only additional information sources about future demand, forecaster should also listen to signals from other functions and search for additional data sources.

As noted in the chapter 3.1.1 target of the forecast should be kept in mind when creating reports. Generally, this is true with analytics. Good foundation is also requirement for efficient analytics and new system often just takes the focus of the problem (see chapter 4.2). Visualization's of the data offer often a good way to understand otherwise complex data. With visualization desirable attributes of the data can be emphasized to easily create value (Heer & Shneiderman 2012). In the case of forecasting visualizations can offer a way to identify trends for example, even if there is not enough data to utilize more complex analytics to identify trends. Visualization are often easy to create and easily understandable. As descriptive analytics this is often enough.

Starting point for the analytics should be the business requirement (see chapter 3.3). Understanding the business requirement, is from the analyst's point of view key skill and the assignment shouldn't be taken as granted (see chapter 3.4). As noted in the introduction, starting point of this research was to improve forecasting capabilities of the target organization. As seen from the Table 5 though, the actual need behind the forecasting depends highly on the function. For example, supply chain wants to improve forecasting to control rising inventory values. In many interviews, it was noted that reason behind rising inventory values is not the inefficient forecast but rather the missing signal of ending project.

As the target organization has hundreds of products only in the Top Hammer button bit – category and much more in the whole product portfolio, understanding the main usages for these is important. Without the understanding of the business valuable insights from data always require someone to validate them and therefore analyst work closely with business and understand it to be able to create value from the data (Davenport & Patil 2012). Target organization offers trainings in their product for their own personnel (see

chapter 4.7). These trainings should be advertised internally, so employees from all the functions could participate and get deeper understanding of the field they work in.

5.3 Suggestion for analytics process for demand forecasting

The main target of this research was to identify analytics process for target organization that supports demand forecasting. Since the scope was forecasting demand of Top Hammer products on individual level, suggested process is especially meant for operative forecasting. As the target organization is currently in the phase of establishing analytics and forecasting, the process should be kept fairly simple. Figure 9 represents developed analytics process for the target organization in the demand forecasting context. It describes the process and relationships of different functions of target organizations during the process. Process is based on the theoretical framework and findings from the interviews.

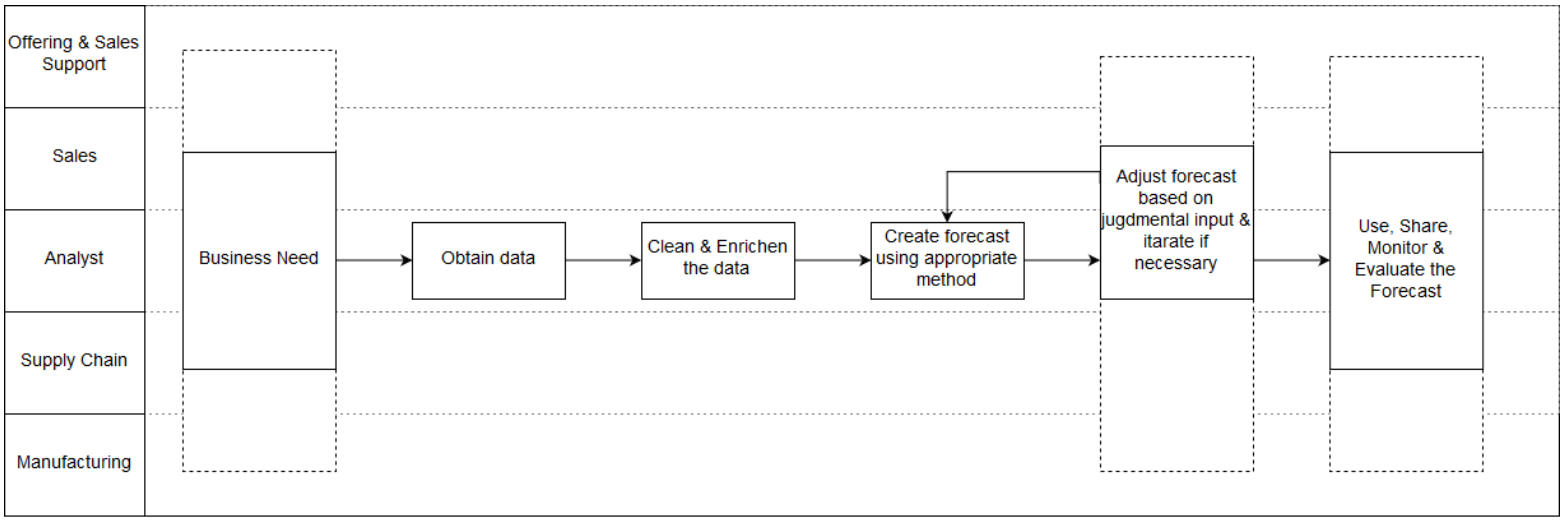

Figure 9. *Suggested demand forecasting process*

Common nominator as a starting point for the analytics in both, theoretical and empirical framework is the business need. By understanding the business need behind the required analytics, analyst must consider the objective of the process and value gained from it. To understand the business need behind the required analytics, evaluation should be done in co-operation of different functions relevant to the case (see chapter 4.5). Differing from BI processes where data is the starting point, starting from the value analyst is not limited to the data currently available. Though, beginning the process from the business need and value might lead to a situation where the answer is that this business need can't currently be met, still valuable insight is gained from the process. This will thrive idea generation in the organization and even if the process fails, information is gained (see chapter 4.8).

In the interviews, main data source for historical data was identified to be the ERP system, NetSuite. Next step in the forecasting should be considering what data is needed to answer the business need. NetSuite offers many attributes which can be included in the reports, but the empirical research noted invoicing to be the best meter for sales. Attributes to be included depend on what is being forecasted. Important attribute to consider is the time

horizon so adequate amount of historical data can be collected. After the required attributes have been selected, report can be run from NetSuite.

Third step of the process is to prepare data for the required format. This includes cleaning the data from unnecessary attributes and possible biases. Initial data quality improvements should also be done by enrichening the data. Data can often be enriched by understanding it and combining it with other data source (Batini et al. 2009). For example, missing information about the product diameter could be enriched by combining product information available in the Robit's public website with the report (Robit Plc).

Next step is the forecast creation. Forecast creation should be done based on the business need. Generally, forecasts can be presented as numerical values, visualizations or some combination of the two. If the business requirement is unclear at this point, input from business should be acquired. The forecasts should still be let to easily editable form, since the fifth step of the process is seeking additional input from other functions about judgmental forecasts and combining qualitative and quantitative information. In the target organization's case it was noted that sales department have generally better view of their customer's future demand than the historical data suggests (see chapter 4.6.1.). Therefore, the judgmental information from the sales function should be highlighted over the historical data. Creating value with analytics is often cycle but forecasting is often linear process (see chapter 3.3). As noted in the chapter 2.1, perfect forecast doesn't exist but iteratively it can be improved. Therefore, during the fifth step of forecasting, iteration should be done until required business needs can be met with the forecasting.

Last step of the analytics process in the forecasting context is using, sharing, monitoring and evaluating the forecast. Value from the forecast should be evaluated based on the original estimated value and whether it was used to answer to the business need. Since the forecast can easily have time horizon of 6 – 12 months, monitoring and updating the forecast should also be done. The last step of the forecasting process in the analytics lasts through the entire time horizon of the forecast.

5.4 Critical evaluation of the research

This research was conducted as embedded case study as defined in the chapter 2.1. When evaluating case studies it is crucial to remember that by definition it is an investigation of a phenomena in specified environment (Saunders et al. 2009; Robson & McCartan 2016). In this research time frame when research was conducted was limited to roughly 12 weeks some of which were during Finnish summer holiday season. Even though researcher had some previous knowledge of the domain, short timeframe set boundaries with number of interviews and gathering of the theoretical framework. These factors should be counted for when reading this research.

First part of this research, review of theoretical frameworks was conducted around two major themes: demand forecasting and business intelligence & analytics. By themselves both are vast subjects that have much literature available around them. This research did not aim to give comprehensive overview of both but rather one point of view on how these areas can be considered in the target organization. Cited literature was selected throughout the research based on Saunders et al. (2009) threefold categorization of literature focusing especially of the secondary level of journals and books. Tertiary literature was used in establishing the fundamentals especially in areas where researcher was not able at first to locate secondary literature. Primary literature which would have been closest to the source with most amount of details was not focused on since the maturity of analytics in the target organization did not enable most cutting-edge methodologies and tools. In addition to these guidelines, amount of citations of the source was emphasized.

Empirical parts of this research have also couple of things to evaluate which affect the quality of this research. It can be said that validity aims to evaluate whether research is valid in its context (Drost 2011) and from that point of view, it is notable that interviewees spoke of the domain quite homogeneously with similar special terminology even though they were chosen throughout the organizational functions. Keeping in mind that interviews were conducted focusing on specific products of the target organization, interviewees can be considered as reliable.

Reader of this research should also consider reliability, the repeatability of a research. As noted above the interview group can be considered reliable and with similar sample research could be repeated. Case study can be said to present the situation as it is in its context (Saunders et al. 2009), which affects how reader should consider reliability of this research. Since the target organization is by their own definition a growth company, context can change swiftly. Therefore, it should be emphasized that this research represents the situation as it was seen in the organization by the interviewees and the literature overview in the time of the research.

Considering the usability of this research, target organization noted results to be useful for them. For a case study this can be seen to improve credibility of the research.

5.5 Next steps

In the empirical part of this research, it was observed, that general maturity of analytics is currently low at Robit. During the interviews repeating themes that rose up regarding development of general analytics and improvements to enhance predictability were noted as follows:

1. Data quality
2. Formalization of external and internal communication
3. Obtaining additional data from dealers about their demand

4. Separating continuous and project demand from the sales data
5. Enhancement of leads and opportunities usage in the NetSuite
6. Extent of the product portfolio in Top Hammer products
7. Optimized production batch sizes

Building the analytical capabilities should be started from the bottom up. Organization needs to understand and develop analytics at the descriptive level before moving to predictive or prescriptive analytics. (see chapter 3.2.4) Upgrading analytics and forecasting capabilities often require investments and evaluating the expected return of the investment (see chapter 2.1). Therefore, in the target organization's case improvement should be started with quick win situations to improve the basic capabilities.

5.5.1 Suggestions for improvements

Improving data quality requires considerable effort, but updating specific missing information is a quick win situation. As noted in the chapter 5.1, there are metadata missing from products details, such as the diameter of the product, in NetSuite. Since this information was noted to be available updating it to products based on their ID good be done swiftly.

Data quality requires naturally iterative updating which in case of the product information should be done through the product's entire lifecycle. Improvements should also be made to the processes of collecting data to ensure data quality will persist when new data is collected. From the forecasting point of view, important data regarding demand type is currently not available. Creating this information for each current sales order and updating this to NetSuite will take time and much effort. Situation should be handled by considering where this kind of data is and how it is collected. As it was noted to be mainly within the sales function, simplest way would be to include demand type information as necessary field to the sales orders creation. This way data about demand type would be collected and quality would increase iteratively.

Theme which affects the predictivity of individual products, repeatedly mentioned in the interviews was the variety of Top Hammer product portfolio. Hundreds of different products which customer can order individually creates problems for many functions. Manufacturing individual products is unfavorable from both subcontracting and own production point of view. Vast product portfolio also requires much training for all stakeholders. As noted in the interviews evaluation of production batches should be done in cooperation with supply chain, production and offering & sales support function at least. Other side of the product portfolio is the consideration of the quantity of different products. From the interviews it was determined that inactivating products within single diameter would be good starting point which would offer clarification in the product portfolio.

Both internal and external communications were noted to have areas that should be improved. Externally, it was noted that additional input from dealers is required to get information about market and future demand. Formalizing this communication, for example by sharing forecasts with dealers would provide additional judgmental data source to the forecasting. Internally more communication especially in formalized way between functions was noted to be required. This would allow Robit to more efficiently identify their internal data sources, not only for forecasting but also for general analytics.

Sales process was presented in the Figure 8. In this, it was noted that offers are made from NetSuite and as noted during the interviews, offers are demand, which either realizes or doesn't. Amount and probability of these offers are currently evaluated in the SOP-meetings which are the most viable source for future sales and projects. Robit's ERP system, NetSuite offers CRM capabilities which would offer more comprehensive handling of the leads and opportunities of upcoming sales.

NetSuite already has these capabilities, but these are not currently used efficiently (see chapter 4.8.1). Enhanced use of these capabilities would require additional project in which sales process would need refinement and training regarding the system would be needed. Requirements for the usage of these CRM capabilities should also be defined so the capability of system could be evaluated. Even though introduction of CRM would allow collection of additional data which would open new data sources for analytics, it would be vast project affecting many functions. Therefore, the project should be started same way as the analytics project: by considering what is the business need behind this.

5.5.2 Future research

Additional research could be done in the future either to improve results of this research or to derive new research. As far as improving results goes, for example additional case studies or action research could be conducted to see how organization with low analytical maturity level could start building up their analytical processes based on this research's analytical process framework.

As the target organization had functions all over the world, international angel offers additional interesting research opportunities. Building analytical capabilities for strongly growing international organization requires systematicity and for example cultural differences should be studied further in order to establish companywide analytical processes. As noted during interviews many customers of the target organization have also quite a different view of the world and would require also additional research in order to establish data flow for the demand forecasting.

REFERENCES

Armstrong, J.S. (2001). Principles of forecasting: a handbook for researchers and practitioners, Springer Science & Business Media,

Batini, C., Cappiello, C., Francalanci, C. & Maurino, A. (2009). Methodologies for data quality assessment and improvement, ACM computing surveys (CSUR), Vol. 41(3), pp. 16.

Bensoussan, B.E. & Fleisher, C.S. (2012). Analysis without paralysis: 12 tools to make better strategic decisions, FT Press,

Bose, R. (2008). Competitive intelligence process and tools for intelligence analysis, Industrial Management & Data Systems, Vol. 108(4), pp. 510-528.

Chen, H., Chiang, R.H. & Storey, V.C. (2012). Business intelligence and analytics: from big data to big impact, MIS quarterly, pp. 1165-1188.

Davenport, T.H. (2006). Competing on analytics, Harvard business review, Vol. 84(1), pp. 98.

Davenport, T.H. & Patil, D.J. (2012). Data scientist, Harvard business review, Vol. 90(5), pp. 70-76.

Drost, E.A. (2011). Validity and reliability in social science research, Education Research and perspectives, Vol. 38(1), pp. 105.

Duan, L. & Da Xu, L. (2012). Business intelligence for enterprise systems: a survey, IEEE Transactions on Industrial Informatics, Vol. 8(3), pp. 679-687.

Díaz, A., Rowshankish, K. & Saleh, T. (2018). Why data culture matters
, McKinsey Quarterly, (4), https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/why-data-culture-matters.

Easterby-Smith, M., Thorpe, R. & Jackson, P.R. (2012). Management research, Sage,

Euchner, J. (2013). What large companies can learn from start-ups, Research Technology Management, Vol. 56(4), pp. 12.

Feigin, G. (2014). Supply Chain Planning and Analytics : The Right Product in the Right Place at the Right Time The Right Product in the Right Place at the Right Time, Business Expert Press, New York,

Fleisher, C.S. & Bensoussan, B.E. (2015). Business and competitive analysis: effective application of new and classic methods, FT Press,

Gilad, B. & Gilad, T. (1985). A systems approach to business intelligence, Business horizons, Vol. 28(5), pp. 65-70.

Hannula, M. & Pirttimäki, V. (2004). A cube of business information,

Heer, J. & Shneiderman, B. (2012). Interactive dynamics for visual analysis, Queue, Vol. 10(2), pp. 30.

Idoine, J. & Herschel, G. (2016). Combine Predictive and Prescriptive Analytics to Drive High-Impact Decisions. (Gartner),

Kaisler, S., Armour, F., Espinosa, J.A. & Money, W. (2013). Big data: Issues and challenges moving forward, IEEE, pp. 995-1004.

Krumeich, J., Jacobi, S., Werth, D. & Loos, P. (2014). Big data analytics for predictive manufacturing control-A case study from process industry, IEEE, pp. 530-537.

Laihonen, H., Hannula, M., Helander, N., Ilvonen, I., Jussila, J., Kukko, M., Kärkkäinen, H., Lönnqvist, A., Myllärniemi, J. & Pekkola, S. (2013). Tietojohtaminen,

Lee, J., Kao, H. & Yang, S. (2014). Service innovation and smart analytics for industry 4.0 and big data environment, Procedia Cirp, Vol. 16 pp. 3-8.

Liebowitz, J. (2006). Strategic intelligence: business intelligence, competitive intelligence, and knowledge management, Auerbach Publications,

Liew, A. (2013). DIKIW: Data, information, knowledge, intelligence, wisdom and their interrelationships, Business Management Dynamics, Vol. 2(10), pp. 49-62.

Loshin, D. (2012). Business intelligence: the savvy manager's guide, Newnes,

Pihlajamaa, T. (2015). Statistical demand forecasting algorithm
, pp. 59.

Robit Plc Robit Plc - public website

Robson, C. & McCartan, K. (2016). Real world research, John Wiley & Sons,

Sallam, R.L., Richardson, J., Hagerty, J. & Hostmann, B. (2011). Magic quadrant for business intelligence platforms, Gartner Group, Stamford, CT,

Saunders, M., Lewis, P. & Thornhill, A. (2009). Research methods for business students, Pearson education,

Schoenherr, T. & Speier-Pero, C. (2015). Data science, predictive analytics, and big data in supply chain management: Current state and future potential, Journal of Business Logistics, Vol. 36(1), pp. 120-132.

Skendzic, A. & Kovacic, B. (2012). Microsoft office 365-cloud in business environment, IEEE, pp. 1434-1439.

Soll, J.B., Milkman, K.L. & Payne, J.W. (2015). Outsmart your own biases, Harvard business review, Vol. 93(5), pp. 64-71.

Stevenson, W. (2014). Operations management, McGraw-Hill Higher Education,

Thierauf, R.J. (2001). Effective business intelligence systems, Greenwood Publishing Group,

Tyrväinen, T. (2013). Business intelligence trends in Finland in 2013,

Vuori, V. (2011). Social media changing the competitive intelligence process: elicitation of employees' competitive knowledge, Tampereen teknillinen yliopisto.Julkaisu-Tampere University of Technology.Publication, Vol. 1001 pp. 2011.

Wang, C., Cheng, H. & Deng, Y. (2018). Using Bayesian belief network and time-series model to conduct prescriptive and predictive analytics for computer industries, Computers & Industrial Engineering, Vol. 115 pp. 486-494.

Wang, G., Gunasekaran, A., Ngai, E.W.T. & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications, International Journal of Production Economics, Vol. 176 pp. 98-110. Available (accessed Cited By :95): https://www.scopus.com/inward/record.uri?eid=2-s2.0-84962360985&doi=10.1016%2fj.ijpe.2016.03.014&partnerID=40&md5=c7f274982b68db381be7bd885ede69c8.

Watson, H.J. & Wixom, B.H. (2007). The current state of business intelligence, Computer, Vol. 40(9),

Witkowski, K. (2017). Internet of things, big data, industry 4.0–Innovative solutions in logistics and supply chains management, Procedia Engineering, Vol. 182 pp. 763-769.